THE PEDAGOGICAL SEMINARY AND

JOURNAL OF GENETIC PSYCHOLOGY

Child Behavior, Animal Behavior, and Comparative Psychology

EDITED BY
CARL MURCHISON

John E. Anderson
University of Minnesota
Charlotte Buhler
Universitat, Wien
William H. Burnham
Clark University
Cyril Burt
University of London
Leonard Carmichael
University of Rochester
Ed. Claparède
Université de Genève
Edmund S. Conklin
Indiana University
Sante De Sanctis
R. Università di Roma
Arnold Gesell
Yale University
William Healy
Judge Baker Foundation, Boston
Leta S. Hollingworth
Teachers College, Columbia University
Buford Johnson
The Johns Hopkins University
Harold E. Jones
University of California
Truman L. Kelley
Harvard University
Yoshihide Kubo
Hiroshima Normal College
K. S. Lashley
Harvard University
A. R. Luria
Medico-biological Institute, Moscow
Toshio Nogami
Kyoto Imperial University
Henri Piéron
Université de Paris
William Stern
Hamburgische Universitat
Calvin P. Stone
Stanford University
Lewis M. Terman
Stanford University
Godfrey Thomson
University of Edinburgh
E. L. Thorndike
Teachers College, Columbia University
C. J. Warden
Columbia University
John B. Watson
New York City
Helen Thompson Woolley
Teachers College, Columbia University

VOLUME 50

1 9 3 7

Copyright, 1937, by The Journal Press
Published quarterly by The Journal Press
Provincetown, Massachusetts, U. S. A.

AUTHOR INDEX

Ayer, Mary Ellen 163
Barry Jr, H 198
Bernreuter, Robert G 163
Bernstone, Arthur H 229
Bousfield, W A. 198
Burtt, Harold E 187
Carmichael, Leonard 217
Carter, H D 157
Chen, Han-Piao 3, 15
Clarke, Frances M. 458
Dove, Claude C 229
Eisenson, Jon 195, 457
Evans, Sol 243
Farnsworth, Paul R 277
Fillmore, Eva A 427
Flemming, Cecile White 323
Forbes, Marie L H 463
Frye, Ellis K 293
Gesell, Arnold 465
Glanville, A Douglass 109
Hansen, Harvey C 361
Hunt, Edna B 458
Hunt, William A 458
Jones, H E 157
Kappauf, William E 27
Kreezer, George 109
Langlie, T A. 339, 401
Lehner, G F. J 217
Lincoln, Hazel 283
Lough, Orpha Maust 307
Maller, J B 91
McGehee, William 441
McGeoch, Grace O 171
Meltzer, H 63
Muenzinger, Karl F 229
Pintner, R 91
Skeels, Harold M 427
Smith, Karl U 137, 371
Schlosberg, Harold 27, 47
Stone, Calvin P. 203, 206
Symonds, Percival M 83
Van Riper, C 193
Warkentin, J. 371
Wilson, Frank T 323
Zigler, Michael J 211

TABLE OF CONTENTS

Sex differences in simple syllogistical judgment . . 3
By Han-Piao Chen

Studies on Thurstone's Reasoning Tests . . 15
By Han-Piao Chen

Conditioned responses in the white rat III Conditioning as a function of the length of the period of delay 27
By William E Kappauf and Harold Schlosberg

Reaction-time apparatus . . . 47
By Harold Schlosberg

Anger adjustments in relation to intelligence and achievement 63
By H Meltzer

Changes in sex differences in problems and interests of adolescents with increasing age . . . 83
By Percival M Symonds

Month of birth and average intelligence among different ethnic groups 91
By R Pintner and J B Maller

A method for the quantitative analysis of human gait . 109
By George Kreezer and A Douglas Glanville

The postoperative effects of removal of the striate cortex upon certain unlearned visually controlled reactions in the cat . 137
By Karl U. Smith

A further study of affective factors in learning . . 157
By H D Carter and H E Jones

A study of the relationship between discipline and personality traits in little children . 165
By Mary Ellen Ayer and Robert G Bernreuter

Reminiscence in maze learning by kindergarten children . 171
By Grace O McGeoch

A further study of early childhood memory . . 187
By Harold E Burtt

Short Articles and Notes 193

- The effect of penalty upon frequency of stuttering spasms (C. Van Riper) 193
- A note on the perseverating tendency in stutterers (Jon Eisenson) 195
- Incidence of orphanhood among fifteen hundred psychotic patients (H Barry, Jr and W A Bousfield) 198

APPARATUS 203

A sand-tube obstruction apparatus (Calvin P. Stone) 203

A paper-window obstruction apparatus (Calvin P. Stone) 206

BOOKS 211

Harvey A. Carr, *An introduction to space perception* (Michael J. Zigler) 211

The development of temperature sensitivity during the fetal period 217
BY LEONARD CARMICHAEL AND G. F. J. LEHNER

Serial Learning II The bi-directional goal gradient in the endless maze 229
BY KARL F. MUENZINGER, CLAUDE C. DOVE, AND ARTHUR H. BERNSTONE

How rats learn the simple alternation problem in a temporal water maze 243
BY SOL EVANS

Continued training with the omission of certain nebenreize 277
BY PAUL R. FARNSWORTH

A study of the Cornell-Coxe performance ability scale with superior children 283
BY HAZEL LINCOLN

The mechanical abilities of siblings 293
BY ELLIS K. FRYE

The effect of functional periodicity on the learning process 307
BY ORPHA MAUST LOUGH

*Correlations of chronological age, mental age, and intelligence quo*tient with other abilities and traits in Grade I 323
BY FRANK T. WILSON AND CECILE WHITE FLEMMING

Personality ratings I Reliability of teachers ratings 339
BY T. A. LANGLIE

Scholastic achievement of Indian pupils 361
BY HARVEY C. HANSEN

The development of visual acuity in the cat 371
BY J. WARKENTIN AND K. U. SMITH

Personality ratings I Validity of teachers ratings in predicting scholastic achievement 401
BY T. A. LANGLIE

The mental development of children from underprivileged homes 427
BY HAROLD M. SKEELS AND EVA A. FILLMORE

The free word association of elementary school children 441
BY WILLIAM MCGEHEE

SHORT ARTICLES AND NOTES . 457

Some characteristics of the written speech of stutterers (Jon Eisenson) 457

Plantar responses in infants following a startle stimulus (Frances M. Clarke, William A. Hunt, and Edna B. Hunt) 458

APPARATUS . . 463

The typewriter key as a device for registering choice. Timing (Marie L. H. Forbes) . 463

BOOKS . . 465

N. Kohts, Infant ape and human child; instincts, emotions, play, habits (Arnold Gesell) . . . 465

BOOKS RECEIVED . 468

The Journal of Psychology

EDITED BY

Carl Murchison

The Journal of Psychology was founded at the request of many psychologists throughout the world for the purpose of establishing a medium in which publication shall be immediate and in which other publication experiments may be carried out. Beginning with the third volume the following methods and terms will be in force:

(1). The size of each volume will be increased to 500-600 pages, and the subscription price will be $7 per volume, two volumes per year.

(2). At the request of the American Library Association, the loose leaf feature will be discontinued, and only the regular library edition will be sent to subscribers. The format of each volume will continue to be such that each subscriber may tear down the assembled volume into its component articles.

(3). Except through the fault of the author, the lag in publication should not be more than three or four weeks.

(4). The author will continue to receive 200 reprints of his paper, and the subscriber will receive the accumulated publications bound into half volume units. The second half of each volume will contain the title page and table of contents for that volume. Each half volume unit will be mailed to the subscriber at its stated date.

(5). The restriction on the length of experimental papers is removed, but theoretical papers must not be more than sixteen pages in length.

(6). Each author will be charged the actual manufacturing costs of any plates or half tone cuts.

(7). The publication charges will be readjusted as follows:

4	page	signature	$10
8	"	"	20
12	"	"	30
16	"	"	40
20	"	"	50
24	"	"	60
28	"	"	70
32	"	"	80
36	"	"	90
40	"	"	100

These charges are tentative and will be adjusted as justified by the subscription list.

(8). *The Editorial Board.* To a much greater extent than is the usual custom the Editorial Board will be drawn from the ranks of the younger psychologists who have shown research promise and have achieved a rank equal to full membership in the American Psychological Association.

(9). The accumulated units will be mailed to subscribers quarterly: January, April, July, and October. The units may vary in size, to accommodate publication needs.

Beginning with January, 1937, one thousand to twelve hundred pages (two volumes) annually. Per annum $14.00; per volume $7.00; per half volume unit $3.50. Complete sets at $7.00 per volume, plus transportation.

THE JOURNAL PRESS
2 Commercial Street
Provincetown, Massachusetts
U. S. A.

$7.00 per volume
Single numbers $4.00

QUARTERLY
Two volumes per year

March, 1937
Volume 50, First Half

Founded by G. Stanley Hall in 1891

THE PEDAGOGICAL SEMINARY AND
JOURNAL OF GENETIC PSYCHOLOGY

Child Behavior, Animal Behavior, and Comparative Psychology

MARCH, 1937

(Manuscripts are printed in the order of final acceptance)

Sex differences in simple syllogistical judgment . . . 3
By Han-Piao Chen

Studies on Thurstone's Reasoning Tests . . . 15
By Han-Piao Chen

Conditioned responses in the white rat. III. Conditioning as a function of the length of the period of delay . . . 27
By William E. Kappauf and Harold Schlosberg

Reaction-time apparatus . . . 47
By Harold Schlosberg

Anger adjustments in relation to intelligence and achievement . . . 63
By H. Meltzer

Changes in sex differences in problems and interests of adolescents with increasing age . . . 83
By Percival M. Symonds

Month of birth and average intelligence among different ethnic groups . . . 91
By R. Pintner and J. B. Maller

A method for the quantitative analysis of human gait . . . 109
By George Kreezer and A. Douglas Glanville

(OVER)

Copyright, 1937, by The Journal Press
Provincetown, Massachusetts
Entered as second-class matter August 3, 1897, at the post-office at Worcester, Mass., under the act of March 3, 1879
Application made for reentry at Provincetown, Massachusetts

The postoperative effects of removal of the striate cortex upon certain *unlearned visually controlled reactions in the cat* 137
By KARL U. SMITH

A further study of affective factors in learning 157
By H. D. CARTER AND H. E. JONES

A study of the relationship between discipline and personality traits in little children 165
By MARY ELLEN AYER AND ROBERT G. BERNREUTER

Reminiscence in maze learning by kindergarten children 171
By GRACE O. MCGEOCH

A further study of early childhood memory 187
By HAROLD E. BURTT

SHORT ARTICLES AND NOTES 193

The effect of penalty upon frequency of stuttering spasms (C. Van Riper) 193

A note on the perseverating tendency in stutterers (Jon Eisenson) 195

Incidence of orphanhood among fifteen hundred psychotic patients (H. Barry, Jr. and W. A. Bousfield) 198

APPARATUS 203

A sand-tube obstruction apparatus (Calvin P. Stone) 203

A paper-window obstruction apparatus (Calvin P. Stone) 206

BOOKS 211

Harvey A. Carr, *An introduction to space perception* (Michael J. Zigler) 211

SEX DIFFERENCES IN SIMPLE SYLLOGISTICAL JUDGMENT*[1]

Tsing Hua University

HAN-PIAO CHEN

INTRODUCTION

This experimental study aims to find whether there is any significant sex difference in simple syllogistical judgment represented by Thurstone's *Reasoning Test, Form B* (8). Special care was taken to control the experimenter-subject sex relationship and to treat the numerical results by rigid statistical criteria

Methodological Considerations It is obvious that the experimenter has some influence on the behavior and consciousness of subjects taking part in any experiment. In fact, this subject-experimenter relationship is itself a psychological problem In the orthodox introspective experiment, where the observer generally has special training, the experimenter has relatively less effect upon the observer's behavior, consciousness, and introspections than in the behavioral experiments, where very little technical training is required of either the experimenter or the subject The experimenter is oftentimes a part of the experimental situation Subjects, especially naive ones, react not only to the particular experimental set-up but also, perhaps primarily, to the experimenter as well It is important that any experiment in sex difference must take account of the sex of the experimenter One pertinent criticism of most researches on sex difference in the past is that they were conducted exclusively by either a male or a female experimenter on subjects or observers of both the same and the opposite sexes

It is well known in hypnosis that *rapport* must be maintained between the hypnotizer and the hypnotized, and, in laboratory experiments on normal subjects, the attitude of the subject toward the experiment itself also affects the results greatly, but very few investigators noticed the importance of the attitude of subjects toward

---

*Received in the Editorial Office on May 28, 1935

[1] This study was directed by Siegen K Chou

the experimenter. It seems to the present writers that in any psychological experiment, the individual experiment especially, the result of a group of female subjects under the conduction of a male experimenter may differ greatly from that under a female experimenter, and vice versa. Since the experimenter is a part of the experimental situation, by which the subject may possibly be stimulated, it is reasonable to believe that the attitude of female subjects toward the male experimenter would be different from that toward the female experimenter. It is hoped that the special precaution of using one male experimenter in this experiment to conduct the test on 27 male subjects and a female experimenter on 27 female subjects may meet this obvious, but neglected, objection in the sex-difference experiment. It must be pointed out, however, that this experimenter-sex control is assumed to be more significant than the avoidance of an equally obvious individual difference in the change of an experimenter.

Reliability of Sex Difference. Bonser in 1910, as reported by Fisher (4), tested 385 boys and 372 girls and found that the median score of the reasoning ability tests of the boys was 165 83 while that of the girls was 159.25 The difference between the sexes was 6.83 in favor of the boys Gates (6) reported his results of investigations on college students In all cases, men excel women about 2.0, 3.1 and 0.8 per cent during the three years, respectively. Lowe and Shimberg (2) had tested 529 males and 600 females and obtained norms for pupils from the 7th to 12th grades. They used Thurstone's *Syllogisms Test,* which was similar but much simpler than the same author's *Reasoning Test, Form A,* used by us in another investigation From Lowe and Shimberg's result, Bronner, Healy, Lowe, and Shimberg concluded that "the difference between the sexes, however, is rather striking In the syllogisms test there is a reliable difference between girls and boys at every grade from seven to twelve in favor of the boys" Brigham (1) had also reached the same conclusion that "the comparison of seventh and eighth grade boys and girls shows slight sex differences in favor of the boys." Recently, Maier (7), in studying human reasoning, also reached the same conclusion that women were markedly inferior to men in problem-solving ability.

On the other hand, Dale concluded, after comparing the test results of boys and girls in the mixed classes taught by women, boys taught by men, and girls taught by women, that "there is no steady sex difference revealed, but merely an indication of the better development of reasoning power in boys taught by men than in girls taught by women" (3) So far as the general results are concerned, it seems that there is some sex difference in reasoning ability. However, is this slight difference in reasoning tests between men and women reliable according to rigid statistical criteria? Bronner, Healy, Lowe, and Shimberg did not give us any concrete data about the so-called "*reliable difference between girls and boys*" (2), nor did the other investigators give us sufficient data for the calculation of the reliability of their differences, hence, it is difficult for us to believe, or not to believe, that the above-reported differences are statistically significant

In the treatment of results reported in this experiment, we applied critical-ratio technique rigidly to our data and have drawn only conservative conclusions for this particular mental performance of syllogistical judgment

The Experiment

Material and Apparatus The Thurstone *Reasoning Test, Form B,* was translated into Chinese for the present experiment of 32 syllogistical arguments, which deal with the simple "tall-short" relationship among three persons, whose actual relative height relationship is different from argument to argument For example, in one argument, Brown may be said to be taller than Jones and shorter than Smith, while, in another, Brown may be shorter than both Jones and Smith, and still, in another, Brown may be taller than both Jones and Smith.[2]

The Alber's Optical Stimulus Apparatus was used for the exposure of the syllogistical arguments, which were written on separate stimulus-cards In front of the apparatus there is a window which can be opened or closed by the experimenter at will Behind that window there is a box which contains all the stimulus-cards When the shutter is opened, the first stimulus-card can be seen by the subject through the window The stimulus-card will be covered up by the metal plate shutter as soon as it is read The next operation opens the window and makes the first stimulus-card fall down to the table, while, at the same time, the next stimulus-card will appear in the window In repeating these operations (opening and closing the window), the stimulus-cards will be shown one by one to the subject in succession

---

[2] Chinese surnames were used in place of the English names

Procedure The experiment was performed under the individual laboratory situation in a small, quiet, well-lighted room. During the test, there were only an experimenter and a subject in the laboratory The subject sat on an adjustable chair, the height of which might be regulated so that the subject felt comfortable and, at the same time, might easily see the window of the exposure apparatus on the table. On the other side of the table, there was a seat for the experimenter This made the experimenter and the subject sit face to face. But between them there was Alber's Optical Stimulus Apparatus, which was sufficiently high to obstruct their eyesight so that neither could see the other's facial expressions. Besides opening and closing the window for the exposure of the different arguments, the experimenter also recorded the time and answer of each judgment Nothing was said by the experimenter about the nature of the experiment After the stimulus-cards had been inserted into the box of the apparatus and the subject's height had been adjusted to the table, the following typewritten English instructions were given

> Please read the following instructions and do exactly what you are told to do You may ask the experimenter any question before the experiment, but do not ask any question when the experiment has already started.
>
> This is an experiment on reasoning It is also a test of concentration You will be shown, through the oblong window in the machine in front of you a series of thirty two arguments in Chinese in the form of syllogisms of the very simple sort Some of these arguments have true conclusions and some have false conclusions Each argument will be exposed through the window as long as you need for reaching a decision as to its truthfulness or falsehood As soon as you reach a decision as to the truthfulness or falsehood of the argument, respond immediately by saying in Chinese *Day* (right) or *Boday* (wrong). Your reasoning response time will be registered by a stop-watch.

> Immediately after reasoning, please write down your introspection about arguments or situations
>
> The experimenter will say "Ready!" and then, after one second, the window of the machine will open to expose the argument for you to read Please work as rapidly as you can without making mistakes

Scoring. Time and error in judging each argument is treated separately in this report An error is counted whenever the subject judged the right conclusion as wrong or the wrong conclusion as right The time in seconds for judging each argument was registered by a stop-watch

Subjects and Experimenters Subjects were secured by personal approach Altogether, we tested 57 subjects on Thurstone *Reasoning Test, Form B* Twenty-seven college men and three college women[3] were tested by a male experimenter, Mr Han-Piao Chen, while the other 27 college women were tested by a female experimenter, Miss Laura Chao, who also made a preliminary report of the results in the senior author's seminar The male experimenter was an assistant in psychology and the female experimenter a junior in psychology The female experimenter served as a subject in the male experimenter's sitting and also received preliminary training for conducting this experiment, hence, she was much benefited by the experiences of the male experimenter

It must be pointed out that the data used in this report were based only on the results of 27 male subjects tested by the male experimenter and 27 female subjects tested by the female experimenter The results of the three female subjects tested by the male experimenter were discarded because of the fact that the text situation for these three female subjects could not be compared with that for other female subjects so far as the subject-experimenter relationship is concerned

THE RESULTS

Speed and Accuracy of Judgment The individual time and error scores are shown in Table 1 The average number of errors for male and female subjects is 4 0 and 5 9, respectively, favoring the

---

[3] The results of these three female subjects were not used in this report

TABLE 1

INDIVIDUAL TIME AND ERROR IN THURSTONE'S *Reasoning Test, Form B,* CONSISTING OF 32 SYLLOGISTICAL ARGUMENTS

Male subjects under male experimenter			Female subjects under female experimenter		
Subjects	Secs per argument	No of errors	Subjects	Secs. per argument	No. of errors
1	12 3	0	28	15 7	5
2	12.6	2	29	17 4	3
3	15.6	3	30	11 4	2
4	9.8	1	31	19 6	1
5	15.9	1	32	15 3	5
6	6 8	2	33	9.1	5
7	18 3	2	34	6 8	1
8	10 2	5	35	8 9	4
9	13 9	4	36	18 6	17
10	9 2	1	37	15 4	0
11	13 1	5	38	16 8	4
12	15 0	10	39	15 0	15
13	11 1	9	40	10 2	5
14	14	6	41	9 2	0
15	19.3	2	42	9 3	15
16	8 5	7	43	13 2	6
17	38 0	7	44	14 8	1
18	12 1	1	45	10.3	0
19	15 8	3	46	9.6	4
20	8 1	5	47	8 9	13
21	14 3	7	48	14 6	4
22	7 8	3	49	10 1	15
23	13 6	2	50	17 8	9
24	27 4	0	51	22 2	2
25	20.2	8	52	9 3	1
26	11 0	9	53	6 3	12
27	21 2	3	54	10 5	9
Total	397 0	108		345 5	158
Mean	14 7	4 0		12 8	5 9
A D	4.5	2 5		3.7	4 3

males, but the average time is 14 7 and 12 8 sec, respectively, favoring the females. The average variation of time score for males is much greater than that for females, and the reverse is true in case of error scores This result agrees with the common belief that women act more immediately and rapidly than men, but men reason better than women. It has not been discovered, so far as the present writer is aware, whether or not women reason quicker than men, but there are several investigators who definitely claim that men excel women in reasoning ability.

Reliability of Sex Difference. Now, let us see whether the difference found in our tests between the two sexes is reliable. Table 2 shows the reliability of mean and variability differences of both time and error scores. It must be emphatically pointed out that none of the differences is statistically reliable, because none of the critical ratios reached the required size, three or greater. Referring to Table 1, it is seen that there is a difference between male and female subjects so far as the variability of time and error scores are concerned. Males are less variable than females in error scores but more variable than females in time score. This fact, in addition to the results of average time and error scores, leads us to think that the slight superiority of error scores for males and that of time scores for females may be due to the difference in variability between them. The critical ratios of the difference in variability for time and error scores are 2.25 and 2.86, respectively, but those of the difference in mean are only 1.28 and 1.65, respectively. Strictly speaking, none of them is statistically significant, but the difference in variability is relatively more reliable in comparison with the difference in mean.

TABLE 2

RELIABILITY OF MEAN AND VARIABILITY DIFFERENCE OF TIME AND ERROR SCORES

	Mean difference							
	Male subjects under male experimenter		Female subjects under female experimenter					
	Mean	σ_M	Mean	σ_M	D_M	σ_{D_M}	$\frac{D}{\sigma_D}$	Per cent chances
Time	14.7	1.25	12.8	.79	1.9	1.48	1.28	89.97
Error	4.0	.56	5.9	1.00	1.9	1.15	1.65	95.05
	Variability difference							
	S.D.	$\sigma_{S.D.}$	S.D.	$\sigma_{S.D.}$	$D_{S.D.}$	$\sigma_{D_{S.D.}}$	$\frac{D}{\sigma_D}$	Per cent chances
Time	6.47	.88	4.13	.56	2.34	1.04	2.25	98.78
Error	2.90	.39	5.22	.71	2.32	.81	2.86	99.79

Why Sex Difference Is Slight. The above results agree with previous findings in that there seems to be a slight sex difference in syllogistical judgment so far as average time and error scores are concerned, but this slight difference is, however, not reliable according to rigid statistical analysis Several explanations for this unreliability of the present test may be offered here

(*a*) *The experimenter-subject sex relationship has been controlled.* It has been pointed out above that we have had a male experimenter conducting the test on 27 male subjects, and a female experimenter, on 27 female subjects It has also been shown that the experimenter is a part of the experimental situation by which the subject may possibly be stimulated. Hence, in conducting an experiment on sex difference of a certain mental ability or capacity, the difference of test score between sexes would be greater if the experiment was conducted by either male or female only. In the present experiment, the experimenter-subject sex relationship has been controlled. The experimental situations are equally favorable to both sexes and, hence, systematic errors that are likely to produce spurious sex differences have perhaps been avoided

(*b*) *Rigid statistical treatment of data lowers the significance of sex difference* From the results of the present test, some sex differences in Thurstone's *Reasoning Test* were revealed in the average time and error scores, but these differences are not reliable according to rigid statistical analysis It is the opinion of the present writer that any conclusion drawn merely from two statistical averages is dangerous Of course, strict statistical analysis may not be valid, but it seems questionable that the slight differences in reasoning ability between sexes reported by many previous investigators are reliable, because most of their conclusions are based on the average scores only

(*c*) *The test is too simple to demonstrate any reliable difference between college men and women* Another explanation for the unreliable slight sex difference in Thurstone's *Reasoning Test* is perhaps that the test is too simple and easy to demonstrate reliably the sex difference in reasoning ability, if there is any. Unless there is really sex difference in syllogistical judgment, further experiments applying Thurstone's *Reasoning Tests* to children or using some other more complex reasoning tests on college men and women should be performed in order to throw more light on this problem

Sex Difference in Nature of Judgment One interesting point brought out by our experiment is that there seems to be a slight sex difference in the nature of judging this true-false test Fritz had repeatedly demonstrated that "there is a marked tendency on the part of the students to give 'true' reactions rather than 'false.' The ratio is approximately 62 per cent to 38 per cent, respectively" (8). In Thurstone's *Reasoning Test, Form B,* half of the 32 arguments are true and the other half are false, which fact was made known to the subject. Wrong judgments are of two types: namely, those mistaking right conclusions as wrong ones (or in Fritz's term the "false" reactions), and those mistaking wrong conclusions as right ones (or in Fritz's term, the "true" reactions). Table 3 shows us clearly that men made more errors when the arguments have true conclusions than when they have false conclusions, while, on the other hand, women made more errors when the arguments have false conclusions than when they have true conclusions. In other words, men made more negative answers or false reactions (63 per cent) than positive answers or true reactions, and women made more positive answers (57 per cent) than negative answers. We do not find any appreciable difference, however, between "true" and "false" reactions when the records of both sexes are combined. Whether there is really a sex difference in preferring to judge an argument to be positive or negative is a problem that is well worth further experimental verification.

TABLE 3

ANALYSIS OF THE NATURE OF ERRORS

	Total errors	Mistaking right conclusions as wrong		Mistaking wrong conclusions as right	
	No.	No.	Per cent	No.	Per cent
Men	108	68	62.96	40	37.04
Women	158	68	43.04	90	56.96
Total	266	136	51.13	130	48.87

SUMMARY AND CONCLUSIONS

We applied individually a Chinese translation of the Thurstone *Reasoning Test, Form B,* to a group of 54 subjects, 27 being college women under a female experimenter and 27 being college men under a male experimenter.

This individual experiment was designed to investigate the much discussed sex difference in simple syllogistical judgment. One significant feature in this particular experiment is the special control of the experimenter-subject relationship. It is pointed out that most investigators have neglected, in the past, this experimenter-sex control. We are convinced that whether a male or a female experimenter is conducting an experiment on sex difference makes a great difference in the results.

Men, on the average, take longer time (14.7 sec.) to pass a syllogistical judgment, but commit less errors (4.0), while women are usually quicker (12.8 sec.) in passing judgments, but commit, on the average, more errors (5.9). However, these differences are not very significant according to rigid statistical treatment. For time, the critical ratio is 1.28, which corresponds to 89.97 per cent chance and, for error, 1.65, corresponding to 95.05 per cent chance. Confirming most researches in the past, sex differences in variability of response are more significant.

There are many reasons for the lack of significant sex differences by our technique of experiment: (1) The experimenter-subject sex relationship has been controlled; (2) rigid statistical treatment of results lowers the significance of any found difference; and (3) Thurstone's test is perhaps too simple to demonstrate any reliable difference between college men and women. However, the fact that more men (63%) tend to mistake right conclusions as wrong and more women (57%) tend to mistake wrong conclusions as right seems to be a characteristic sex difference.

REFERENCES

1. BRIGHAM, C. C. The diagnostic value of some mental tests. *Psychol. Monog.*, 1917, **24**. Pp. 254.
2. BRONNER, A. F., HEALY, W., LOWE, G. M., & SHIMBERG, M. E. A manual of individual tests and testing. Boston: Little, Brown, 1927. Pp. x+287.
3. DALE, A. B. A group test of reasoning ability. *Brit. J. Psychol.*, 1926, **16**, 314-338.
4. FISHER, S. C. Arithmetic and reasoning in children. *Ped. Sem.*, 1912, **19**, 48-77.
5. FRITZ, M. F. Guessing in a true-false test. *J. Educ. Psychol.*, 1927, **18**, 558-561.
6. GATES, A. I. Experiment on the relative efficiency of men and women in memory and reasoning. *Psychol. Rev.*, 1927, **24**, 139-146.

7 Maier, N. R F An aspect of human reasoning. *Brit J Psychol.*, 1933, **24**, 144-155

8. Thurstone, L L Thurstone reasoning test, Form B Chicago Stoelting

Tsing Hua University
Peiping, China

STUDIES ON THURSTONE'S REASONING TESTS*[1]

Tsing Hua University

HAN-PIAO CHEN

INTRODUCTION

So far as the writer is aware, there seems to be no formal report on the two forms of the Thurstone *Reasoning Test*. It is interesting to know how far these simple syllogism tests can actually test reasoning ability It has been criticized by many people as too simple to be an adequate test for the complex reasoning processes. However, its very simplicity, the familiarity of the test material, and the uniform representation of the materials in the test may be an advantage over those complex, variable reasoning ability tests which more or less suffer from what had been found by Thorndike (3) and Wilkins (4) concerning the effect of changed data and material upon reasoning The aim of the present investigation is not so much a study on reasoning ability as an experimental analysis of Thurstone's *Reasoning Tests* themselves The following problems were in mind when we worked on this test

1 What are the most difficult arguments in these obviously simple, easy reasoning tests? Is there an effect of form and content on the score of judgment?

2 What are the methods of attack in solving these arguments? Do the subjects spontaneously arrive at a certain method of attack at the beginning of the test without trial and error?

3. Is there any correlation between error and time in judging? In other words, does the quicker subject tend to produce more or less valid answers?

4 Is there any correlation between the score of Thurstone's *Reasoning Tests* and the grade in general physics of college students?

EXPERIMENTAL PROCEDURE

Material. The Thurstone *Reasoning Tests, Forms A* and *B,* were

---

*Received in the Editorial Office on May 28, 1935

[1]This study was directed by Siegen K Chou

translated into Chinese for the following four series of experiments. *Form A* consists of 20 syllogistical arguments, which deal with the light-heavy relationship among silver, copper, and iron, the relations among relatives, the relationships of debts among three persons, the relation of club membership between two clubs, and the birthday relationship among three persons. Each of these five types of problems consists of four arguments. *Form B* consists of 32 syllogistical arguments, which deal with the simple tall-short relationship among three persons, whose actual relative height relationship is different from argument to argument For example, in one argument, Brown may be said to be taller than Jones and shorter than Smith, while, in another, Brown may be shorter than both Jones and Smith, and still, in another, Brown may be taller than both Jones and Smith All the arguments, in either *Form A* or *Form B*, are similar in form but none of them is exactly the same in statement.

Test Series. Four series of tests were performed Series I was performed under the individual laboratory situation. Fifty-seven subjects had been tested by means of the Thurstone *Reasoning Test, Form B*. The Alber's Optical Stimulus Apparatus had been used for the exposure of different arguments. The subject was required to read one argument at one time and then judge its truthfulness or falsehood Both time and error for each argument had been recorded by the experimenter separately (1). Series II was performed before the regular hour of a seminar. Ten subjects, sitting around a table in a small library of the Psychology Department, had been tested by *Form A* only (in English). Series III and IV were performed in two lecture rooms The former was given at the regular hour of a class of advanced experimental psychology, while the latter was given at the regular hour of a class of general physics In both series, subjects, sitting in pairs, had been tested by *Forms A* and *B*, which were printed in Chinese on two separate sheets of paper. In order to avoid copying the judgments of one's neighbor, *Forms A* and *B* had been distributed alternatively. This is to say, in the first trial, all subjects whose seats were odd numbers were tested with *Form A* and subjects whose seats were even numbers were tested with *Form B*, and, in the second trial, the forms were reversed. Eight subjects had been tested in Series III, but only seven of them had completed their tasks. In Series IV, 170 students had been tested, but only 155 and

141 of them had answered all the arguments in *Forms A* and *B*, respectively Hence, we had secured 172 subjects for *Form A* and 205 subjects for *Form B*. All series, except I, were tested under group situations The task of the subject in these group series was the same as that in individual series Oral instructions were given before the test by the experimenter After all the subjects had understood what the experimenter wanted them to do, test materials were distributed to them. At a given signal the test was begun The total time required to finish the whole test was recorded by the subject himself immediately after he had finished

AN ANALYSIS OF THURSTONE REASONING TESTS

Relative Difficulty of Different Arguments The relative difficulty of different arguments in *Form A* and *Form B* in terms of error score by all the subjects and *Form B* in terms of time score by 57 subjects are shown in Table 1 This table shows that different arguments are of very unequal difficulty so far as both the time and error scores are concerned Thus, we see that argument No. 17 of *Form A* has only 4 errors but No. 14 has as many as 80 errors, and No. 13 of *Form B* has only 9 errors but No 5 has as many as 49 errors. The time scores show smaller variation. It was the subjects' introspective records of *Form B* in Series I and the experience of the writers during the translation and preparation of this test that suggested that we analyze and compare the different arguments as to their relative difficulty. Several tentative conclusions may now be drawn as follows:

(*a*) *The difficulty of arguments in terms of error score vary with the content, data, or category of arguments.* It has already been pointed out that in *Form A* we have five types of problems to be reasoned out by our subjects These five kinds of relationships are what we mean by different data or content in the tested arguments. It is noted that the total number of errors are different for each kind of relationship The debts relationship, which yielded only 53 errors, is the easiest, and the club membership relation, which yielded 188 errors, are the most difficult to judge The other three categories, birthday, light-heavy, and relatives, yielded 60, 74, and 84 errors, respectively This effect of content upon the relative difficulty of arguments confirms the general findings of Thorndike (3) and Wilkins (4)

TABLE 1
RELATIVE DIFFICULTY OF DIFFERENT ARGUMENTS

Argument No	Total error scores Form A (172 subjects)	Total error scores Form B (205 subjects)	Time score Form B (57 subjects)
1	23	41	19.8
2	46	32	16.5
3	21	34	20.8
4	38	16	15.1
5	8	49	18.2
6	21	33	18.5
7	19	20	16.1
8	16	28	14.6
9	60	15	10.3
10	18	12	11.0
11	19	30	12.2
12	15	12	14.1
13	8	9	11.6
14	80	20	12.8
15	15	18	15.5
16	11	32	12.7
17	4	30	14.4
18	8	17	11.4
19	10	17	12.3
20	19	23	13.2
21		21	11.4
22		36	14.2
23		13	11.1
24		31	12.4
25		22	12.1
26		27	11.3
27		23	11.4
28		26	10.0
29		14	11.0
30		28	10.1
31		10	11.8
32		23	12.0
Total	439	762	432.9

(*b*) *Given the same data or content, the difficulty of arguments in terms of both time and error scores may vary with the tall-short relationship in the argument.* In analyzing *Form B,* it is noted that the 32 arguments may be classified into four groups, or eight sub-groups, according to the various tall-short relationships. The argument, e.g., *Brown is shorter than Smith, Jones is shorter than Brown,* therefore, *Jones is shorter than Smith,* consists of three propositions consistently dealing with the relationship of *shortness,* which may be denoted by the symbols *SSS*. On the other hand, if all three propositions deal with the relationship of *tallness*, then the symbols will be *TTT*. Similarly, we have arguments of *TTS, SST, STT, TSS, TST*, and *STS*. The relative difficulty of the different groups in terms of time and error scores are shown in Table 2.

TABLE 2

Effect of Changed Tall-Short Relationship of *Form B* upon Difficulty

	Group A		Group B		Group C		Group D	
	TTT	*SSS*	*TTS*	*SST*	*STT*	*TSS*	*TST*	*STS*
Time	11.8	11.9	12.5	13.6	13.6	14.4	14.3	16.2
Av	11.9		13.0		14.0		15.3	
Error	57	105	100	97	93	90	96	124
Av	162		197		183		220	

It is easily seen from this table that the arguments under *Group A* are much less difficult than *Group D* by both time and error scores. *Group C*, having 183 errors, seems a little easier than *Group B*, having 197 errors, but the reverse is true in time scores.

This effect of changed relationship upon the relative difficulty of different arguments is by no means an unexpected phenomenon. According to introspective comments, it is to be noted that many subjects frequently reported that those arguments consistently dealing with the relationship of shortness or tallness are easy to comprehend. Moreover, some objects reported that they could not reach a decision unless they had transformed an argument into a *TTT* or *SSS* type. This is to say that some of our subjects used tallness or shortness as the standard of comparison. Whenever he used tallness as his standard, he would change those propositions dealing with the relationship of shortness into tallness, or vice versa. For example, the proposition "Brown is shorter than Smith" would be changed into "Smith is taller than Brown," if the subject used tallness as the standard of reference. The relationship of *TTT* and *SSS* is perhaps easier for memorizing than any other relationship, and so they tend to make the subject have quicker decision and less error.

(*c*) *Given the same data or content, the difficulty of arguments in terms of both time and error scores vary with the sequence of persons in the arguments.* The sequence of persons is different in the arguments of *Form B*. Generally, they may also be classified into four groups of eight types as follows.

Group 1. Type A *(a-b, a-c; c-b)* Type B *(a-b; a-c, b-c)*
Group 2: Type C *(a-b, b-c, c-a)* Type D *(a-b, b-c, a-c.)*
Group 3: Type E *(a-b; c-a, b-c.)* Type F *(a-b; c-a, c-b)*
Group 4: Type G *(a-b, c-b; c-a.)* Type H *(a-b; c-b; a-c)*

Here our interest is only in the sequence of three persons in different arguments. No attempt has been made to analyze the effect of a different sequence of particular persons upon the difficulty of arguments. Each of the three symbols may represent each of the three persons, but each of them represents one person consistently only in that particular argument. For example, *a-b* may represent "Brown is taller (or shorter) than Jones," "Brown is shorter (or taller) than Smith," "Jones is taller (or shorter) than Smith," or any other proposition. Likewise, *a-c, b-c, c-a,* or *c-b* may represent any one proposition in the test; and also *a-b, a-c, b-c* may represent Argument No. 1 (Brown is taller than Jones, Brown is shorter than Smith, therefore, Jones is taller than Smith), or argument No. 6 (Smith is shorter than Jones, Smith is taller than Brown, therefore, Jones is shorter than Brown). The symbol *a* may represent "Brown" in Argument No. 1 but may also represent "Smith" in Argument No 6, although it can represent only one person consistently within any one argument.

The time scores of 57 subjects and the error scores of 205 subjects are shown in Table 3. It is noted that the order of difficulty

TABLE 3

EFFECT OF CHANGED SEQUENCE OF PERSONS UPON DIFFICULTY OF *Form B*

	Group 1		Group 2		Group 3		Group 4	
	A	B	C	D	E	F	G	H
Time	13.9	16.7	11.6	11.6	14.7	11.9	15.5	12.3
Av.	15.3		11.6		13.3		13.9	
Error	116	140	110	98	89	62	82	65
Av.	256		208		151		147	

according to the time score is: Groups 1, 4, 3 and 2, while the error score is: Groups 1, 2, 3, and 4. Although arguments in Group 2 yield more errors, so far as the most difficult arguments are concerned we have reason to believe that those arguments in Group 1 are the most difficult in both time and error scores.

From the results of this and the preceding sections, it seems that the most difficult arguments are those which belong to Group D of the tall-short relationship and, at the same time, belong to Group 1 of the sequence of persons. Table 4 gives this kind of analysis, where A, B, C, and D denote different classes of height relationship, and 1, 2, 3, and 4 denote those groups of sequence of persons, and also A-2, A-3, etc., denote those arguments that belong to both Group A class of height relationship and Groups 2, 3, etc., of sequence of persons. This combined effect of height relationship and sequence of persons is shown in Table 4.

TABLE 4

RELATIVE DIFFICULTY OF ARGUMENTS OF *Form B*

	Combination							
	A-2	A-3	B-2	B-3	C-1	C-4	D-1	D-4
Time	11.1	12.7	12.2	13.8	13.4	14.5	17.2	13.3
Error	95	67	113	84	103	80	153	67

The range of time and error scores of height relationship and sequence of persons taken together are all greater than either taken alone. In other words, the easiest and the most difficult arguments are usually those combining both the easiest and the most difficult height relationship and sequence of persons as seen in Table 5. The most difficult arguments are those combining both the most difficult height relationship and, at the same time, the most difficult sequence of persons. For example, D is found to be the most difficult height relationship and 1 the most difficult sequence of persons by both

TABLE 5

COMPARISON OF RANGE OF TIME AND ERROR SCORES OF *Form B*

	Height relationship	Sequence of persons	After combination
Time	11.8—16.2 (A)—(D)	11.6—16.7 (2)—(1)	11.1—17.2 (A-2)—(D-1)
Error	57—124 (A)—(D)	62—140 (3)—(1)	67—153 (A-3)—(D-1) (D-4)

time and error scores, and hence the combination of D-1 is found to be the most difficult as shown in Table 4

Method of Attack in Solving Problems. From both subjects' introspective records and experimenters' observations during the performance of the test, several methods used by our subjects in judging whether the conclusion of the argument is true or false have been noted. Some had chosen three fingers, two feet and one hand, or three digits to represent the three persons in the argument, and, at the same time, the relative position of the right and left, front and back, or top and bottom was taken as an indication of the tall-short relationship of the persons in the argument Some also used three lines of unequal length to represent three persons, the longest line standing for the tallest person and the shortest line the shortest person One subject reported that she imagined the tallest person located in the highest part of the brain and the shortest person in the lowest part of it Several subjects used the symbols > and < to represent the tall-short relationship. Some other subjects read out the arguments in order to obtain a decision for the truth or falsehood of the consclusions Another interesting fact is that one subject reported that, when he read the argument, he pronounced the name of different perons with different pitches in order to arouse different auditory imagery. A forceful pronunciation of the name of a person means tallness of that person, and a weak pronunciation means shortness.

Those who formulated any method of attack at all acquired it only after some preliminary trial and error. None of the subjects, according to their introspective records, developed any definite method during the first few trials Usually two or three cues were tried out at first by most of the subjects, and then they finally arrived at one definite method of attack

Correlation between Time and Error Scores There are two ways of correlating time with error scores, one with respect to arguments and the other with respect to subjects With respect to arguments, the coefficient of correlation by the product-moment method is found to be +.71; with respect to subjects, —.002 The former correlation indicates that those arguments that required a shorter time for our subjects to judge also yield less error, and vice versa. This result agrees with the data shown in Tables 2, 3, and 4 and indicates that, even thought the arguments are very simple, there is a wide range of relative difficulty among them, judging from both time and error scores.

The time and error correlation with respect to subjects is only — 002, which is practically zero correlation This means that some subjects may take a shorter time and have less error, other subjects may take a longer time but have more errors, still others may take a shorter time but have more errors, and still others may take a longer time but have less errors This tendency may be seen in Table 6 Subjects whose time scores fall within the first quartile, third quartile, or between first and third quartiles are said to have quicker speed, slower speed, and normal speed, respectively Similarly, subjects may be said to have less error, more error, and moderate error It is easily noted from Table 6 that a subject may have slower speed and more error, quicker speed and more error, normal speed and moderate error, and so forth. The two extremes represent those whose time and error scores are positively and negatively correlated, respectively, but the majority of them belong to the class of normal speed and moderate error This result confirms Dale's finding that "there is no relationship at all between score (or number of errors) and rate of working" (2) in reasoning ability tests"

TABLE 6

RELATION BETWEEN TIME AND ERROR SCORES OF *Form B*

(57 subjects)

		Speed			
		Slower	Normal	Quicker	Total
Error	Less	3	6	3	12
	Moderate	8	16	8	32
	More	3	7	3	13
Total		14	29	14	57

Correlation between Scores of Thurstone's Reasoning Tests and Grade in General Physics Incidentally, it may be of interest to know how Thurstone's *Reasoning Tests* are related to the actual grade in the general physics course Before we present the correlation coefficients between test scores and grade marks in this course, a few words should be said about the marking system as used in Tsing Hua University, which has five-letter grades *E* for excellent, *S* for superior, *N* for normal, *I* for inferior, and *F* for failure. The percentage equivalents are as follows: *E* = 91-100 per cent, *S* = 81-90 per cent, *N* = 71-80 per cent, *I* = 61-70 per cent, and *F* = 60 per cent and below. Besides these gross marks, plus or minus is also attached to each grade, and hence twelve marks are possible for the four passing letter grades, each having a value of three points. A grade below *I* cannot be included in the correlation table because it ranges anywhere from zero to 60 points.

We calculated 12 correlation coefficients for both forms of the test, for both time and error scores, and for first and second terms and for the whole academic year. The number of persons varies somewhat in the different correlations. For test *Form A*, 120 persons have both test scores and grades above *I* for the first term, 121 persons for the second term, and 118 persons for the final; for test *Form B*, 107 persons have both test scores and grades above *I* for the first term, 106 persons for the second term, and 103 persons for the final. The correlation coefficients are shown in Table 7.

TABLE 7
CORRELATION BETWEEN THE SCORES OF THURSTONE'S REASONING TESTS AND THE GRADE IN COLLEGE PHYSICS

Test	Score	1st term	2nd term	Final	Average
Form A	Time	—375	—557	—700	—544
Form A	Error	—779	+.070	—421	—377
Form B	Time	—735	—690	—673	—699
Form B	Error	—386	—.738	—649	—591
Average		—569	—479	—611	—553

Since both time and error scores are negative measures of the test trait, it is obvious that all except one of the correlation coefficients in Table 7 are minus. We are not in a position to explain the wide discrepancies in the 12 correlation coefficients. The average in the first term for both tests according to both time and error scores is somewhat higher than that in the second term. When the final grade is correlated with the test, the coefficient is highest. In general, *Test B* seems to correlate higher with grades in physics than *Test A*, and time scores seem to correlate higher than error scores.

Summary and Conclusions

We applied Thurstone's *Reasoning Test, Form A* (English), to a group of 10 subjects and a Chinese translation of both *Forms A* and *B* to 162 and 148 subjects, respectively, under two group situations Another group of 57 subjects had been tested by *Form B* only under individual laboratory conditions The time required to answer each argument was registered in the individual test situation, but only the total time required to complete each form of tests had been recorded in the group test situation

A detailed analysis of Thurstone's *Reasoning Test* revealed the fact (1) that the relative difficulty of the different arguments in terms of error score may vary with the content, data, or category of arguments, (2) that, given the same data or content, the difficulty of arguments in terms of both time and error scores may vary with the tall-short relationship in the argument, and (3) that, given the same data or content, the difficulty of arguments in terms of both time and error scores may vary with the sequence of persons in the arguments. Subjects used all sorts of imaginal cues in passing a judgment, and there is practically no correlation between speed and accuracy of judgment Finally, we found that test scores in Thurstone's *Reasoning Tests* are correlated with grades in college physics to the extent of —611 Since the scores are time and errors, the correlation coefficient is necessarily minus

References

1 Chen, H -P Sex differences in simple syllogistical judgment *J Genet. Psychol*, 1937, **50**, 3-13

2 Dale, A B Group tests of reasoning ability *Brit J Psychol*, 1926, **16**, 314-338.

3 Thorndike, E L The effect of changed data upon reasoning *J Exper Psychol*, 1922, **5**, 33-38

4 Wilkins, M. C The effect of changed material on ability to do formal syllogistic reasoning *Arch of Psychol*, 1928, No 102

Tsing Hua University
Peiping, China

CONDITIONED RESPONSES IN THE WHITE RAT. III CONDITIONING AS A FUNCTION OF THE LENGTH OF THE PERIOD OF DELAY*

University of Rochester and Brown University

WILLIAM E KAPPAUF AND HAROLD SCHLOSBERG

INTRODUCTION

The formation of conditioned reactions in the white rat based upon the unconditioned response to shock has been described in previous papers by Schlosberg (9, 10) He found it quite easy to establish conditioned breathing reactions to a buzz or light but rather difficult to obtain a stable withdrawal of the leg or tail. The influence of certain factors in the experimental situation upon the progress of the learning was investigated, but no attempt was made to vary the interval between the conditioned and unconditioned stimuli The present study was designed to determine the effect of this particular variable

In the early work of Pavlov (6) on "simultaneous," "delay," and "trace" reflexes, time intervals were recorded in minutes rather than seconds But when one turns from a study of glandular responses to one of skeletal activity, precise temporal relationships seem to be more important Thus, Schlosberg (8), using a bell-tap sequence in establishing a conditioned knee-jerk in human subjects, observed that intervals between 2 and 4+ sec led to optimal conditioning, while shorter periods of "delay" were much less favorable Comparable results were reported by Hilgard (2) for the conditioned eyelid reflex More elaborate than the above have been the recent investigations of Wolfle (12, 13). In the course of her two studies of the conditioned hand withdrawal, she employed some 160 subjects, assigning them to different groups and training them with different intervals between the associated stimuli The positive intervals between 2 and 6 sec were most effective, and on either side of this region the incidence of conditioned responses during the test trials was much lower than the maximum frequency (37 per cent in the

---

*Received in the Editorial Office on April 24, 1936

first experiment and 57 per cent in the second) Razran (7) has suggested that the diminished effectiveness of the longer forward intervals employed by Wolfle might be a function of the relatively short period between successive paired stimulations. It may also be mentioned that Wolfle adopted a discrete conditioned stimulus, namely, the impact of a sound hammer Her situation was typically that of a "trace" rather than a "delay" reflex, and this fact also may have affected conditioning at the longer intervals.

The first approach to the problem of optimal temporal sequences in animal conditioning is found in the work of Carr and Freeman in 1919 (1) In this experiment with the white rat in a double-choice maze, the sound of a buzzer was associated with a closed door which made necessary a series of turning and retracing movements Learning to turn at the presentation of the sound was more successful when the buzzer was sounded approximately 1 sec. before the animal reached the door than it was if the two stimuli were coincident or if the buzz came some 8 sec after the retracing had begun. In a very similar situation, Yarborough (14) established the habit of turning and retreating from the pathway when a shock was received Subsequent presentations of a buzz in definite time relations with the shock formed the basis for a first-order conditioned response. With this technique, intervals greater than 2 sec were regularly less satisfactory, while those of 1 and 0 sec were both more effective in bringing the animals to the 90 per cent criterion. When second-order conditioning was attempted with a new sequence of light and buzz, the order of difficulty increased with the length of the interval

Perhaps a more typical conditioning situation was that devised by Warner (11). The unconditioned response was the rat's uncoordinated scrambling over a low wall which separated a charged from an uncharged grid. The conditioned stimulus, a buzz of 1-sec. duration, preceded the shock by 1, 10, 20, or 30 sec for different experimental groups. Of these groups, the one trained with the 1-sec delay reached the criterion most rapidly

This brief account covers the literature on controlled investigations of the relative effect of different short intervals upon the establishment of conditioned responses so far as the present writers are aware. Other experiments seem to have adopted intervals which were either traditional, convenient, or which afforded satisfactory results This latter is clearly the case in those studies of auditory sensitivity involving the conditioned breathing reaction A 5- or 10-sec. interval of delay affords a breathing record adequate for analysis, but there is no reason to believe that this is the most satisfactory period for the development of sharp, conditioned skeletal responses In the earlier work of Schlosberg on the white rat, the period of delay was set at 1/3 sec to correspond approximately to the optimal interval for humans In contrast to the reported consistency of conditioned skeletal reactions in the higher mammals (5), the incidence of leg or tail responses in the rat was relatively low. It seemed quite possible that the conditioning might be more effective with an interval other than 1/3 sec

In investigating the importance of this temporal factor, the present experiment has given no consideration to "backward" intervals because of their relative ineffectiveness (6, 12, 13)

APPARATUS AND GENERAL PROCEDURE

Most of the parts of the apparatus used in the present study have already been described (10) The holder was so arranged that the rat was supported astride a block of wood. His right foreleg was held extended by the tension of a light mechanical recording system; the other three legs were secured by thongs A second lever arrangement made it possible to record tail movements, and a pneumatic system (balloon and tambour) provided a record of respiratory activity The unconditioned stimulus was a 60-cycle half-wave rectified alternating current, applied through a wick electrode on the right foreleg and a large "indifferent" electrode on the abdomen A resistance of 5 meg was put in series with the animal, and the current, as read on the milliameter, was controlled by varying the voltage A loud buzzer furnished the conditioned stimulus

In place of the original timer, it was necessary to substitute one which would permit a wide variation of stimulus durations For this purpose a Leipsig time-sense apparatus was rebuilt and driven by a Telechron motor. Four trip keys on the graduated circumference were connected to appropriate relays and switches so that the successive opening of these keys would initiate the buzz, then the shock, and finally terminate both simultaneously. The temporal relations between these events were regularly checked with an impulse counter or an electric stop-clock, which could be read to 1/120 or to 1/300 sec, respectively.

After a consideration of the previous conclusions of Schlosberg (9, 10), Liddell (5), and others, the following training technique, which seemed most favorable to conditioning, was adopted. From 20 to 25 buzz-shock sequences were presented during each of six daily sessions The interval between successive paired stimulations was never less than one minute and rarely more than two, so that each session lasted about 35 minutes The stimuli were never deliberately given while the animal was struggling. Since it has often been suggested that unreinforced "test" trials delay the establishment of the conditioned response, the buzz was never presented without the shock In consequence, the sole criteria of the progress of conditioning were early responses, i e, behavioral changes which appeared after the buzz had begun but before the shock was administered

The time intervals which were selected for special study were 1/3, 2/3, 1, 2, 4, and 7 sec Following these fixed periods of "delay," the buzz and shock continued together for 1/3 sec.

The animals used were healthy, well-fed males descended from Wistar Institute Commercial stock. They had had no previous training with buzzer, shock, or restraint conditions, and were taken at random from that group in the colony which happened to be between five and eight months of age at the time of the experiment Two animals were always trained concurrently at a given delay period and were never used for more than one such interval. In each experimental pair, one animal was trained with a shock of fixed duration The shock circuit for the other animal included a short-circuiting key adjusted so that any leg withdrawal exceeding 1 mm, if maintained until the current flow had ceased, would prevent or terminate the shock Schlosberg (9, 10) has already reported that there seems to be little difference in effectiveness between these two methods even though the *SC* animal consistenly terminates the shock by his response to it. The two situations were carried over into the present study for further analysis Thus, one rat of each pair received a shock of 1/3-sec. duration, while the other animal usually "shorted" the shock at the end of .1 sec

EVALUATION OF THE RESPONSES

In general, the records of the present experiment add further support to the statement that the conditioned response does not precisely reproduce the unconditioned response in form or magnitude (2, 8, 9, 10). When the reactions were originally classified, particular attention was given to those leg flexions, breathing changes, and movements of the tail which approximately corresponded in suddenness and amplitude to the unconditioned activity elicited by shock But such breathing reactions never appeared in more than 30 per cent[1] of the trials for any one animal, and, on the average, occurred in less than 12 per cent Complete leg flexions had an average incidence of less than 8 per cent, and vigorous tail movements were still more infrequent It was, therefore, very impractical to try to use these responses as indices of the depth of conditioning In the final data sheets and in the accompanying charts, all leg flexions were grouped together and all whipping and bending movements of the tail were amassed regardless of their relative magnitudes Besides flexion of the members, relaxations and extension were also observed Very frequently the limb was actively thrust forward or the tail was stiffened so that the markers made excursions in directions opposite to those denoting flexions Such responses definitely involve different muscle systems and hence were grouped apart from the flexions which resembled the unconditioned reaction more closely.

Respiratory changes were classified under several headings.

1. All sharp inspirations which were of an amplitude at least three times that of the previous breathing pattern
2. "Regular" squealing or whispering, i e, the continued repetition of the squeal pattern until the presentation of the shock When not vocalized, the forced inspiration could often be heard as a "whisper"
3. All other modifications in rate, regularity, etc.

Simple squealing always occurs when a shock is administered, but it fails to be a good criterion of conditioning because of the fact that it attends expiration rather than inspiration At the shorter intervals, an animal may receive the shock before the conditioned inspiration is complete. Thus squealing could be of value only at the longer intervals. As a part of the conditioned reaction, it will be discussed in a later portion of the paper, but it was impossible to quantify squealing as a measure for graphical representation Finally, since slight changes in the breathing pattern might have been mere sensitizations of the original unconditioned reaction to buzz (which was extinguished on the first day), the writers believe that the incidence of sudden inspirations of a given magnitude provides one of the best measures of conditioned respiration

---

[1] In only two animals did they appear in more than 20 per cent of the trials

To avoid the difficulties of attempting to evaluate responses which were masked by struggling or squealing which began before the onset of the buzz, all such trials, although contributing to the training, were discarded in the final scoring. All the records were read and graded at least twice.

METHODS AND RESULTS OF THE MAIN EXPERIMENTAL GROUP (*A*)

The first experimental group of 12 animals was trained under "standard" conditions to determine the relative effectiveness of the different intervals which had been chosen. Two animals of the same age and stock, but not necessarily the same litter, were used at each of the six chosen intervals, and for one member of each pair the short-circuiting contacts were introduced Throughout the six sessions, the shock strength was kept at the constant level of 2 ma All factors except the length of the delay period and the presence of the short-circuiting contacts were identical for every animal of the group. In order to throw some light upon the significance of the observed differences between the "short-circuit" and the "non-short-circuit" groups and between the performance displayed at one interval and at the next, four additional animals were run as controls. One pair was trained at 1/3 sec., the other at 2/3 sec The short-circuiting contacts were not used with these control animals.

Because the curves of daily performance were characterized by marked irregularities and were not very suggestive, the behavior of these 16 subjects was summarized in terms of the percentages of trials in which significant responses appeared. This percentage incidence was computed for the entire six-day training period, and again for the last three days The same trends are to be observed

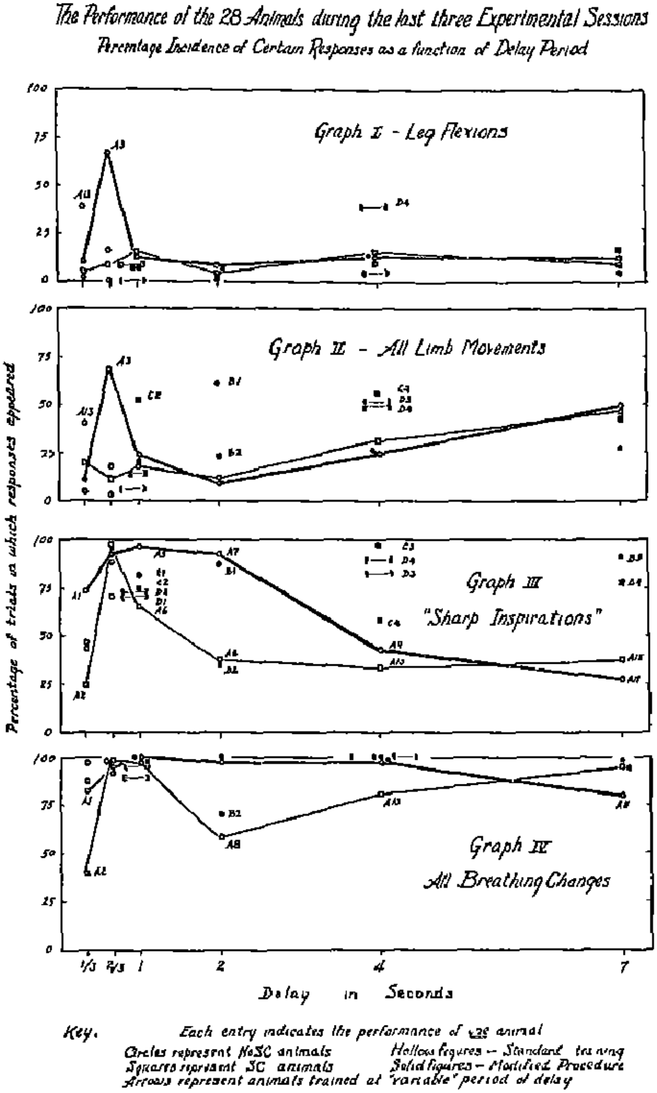

The Performance of the 28 Animals during the last three Experimental Sessions
Percentage Incidence of Certain Responses as a function of Delay Period
Graph I - Leg Flexions
Graph II - All Limb Movements
Graph III "Sharp Inspirations"
Graph IV All Breathing Changes
Percentage of trials in which responses appeared
0
25
50
75
100
1/3
2/3
1
2
4
7
Delay in Seconds
Key.
Each entry indicates the performance of one animal
Circles represent NoSC animals
Squares represent SC animals
Arrows represent animals trained at "variable" period of delay
Hollow figures — Standard training
Solid figures — Modified Procedure

FIGURE 1

in the two sets of data so procured. Since the reader may prefer an account of the final level of performance rather than a measure which masses all phases of the training, the accompanying charts (Figure 1) summarize the incidence of *early* responses in the *three concluding sessions*. In considering these data it is important to bear in mind that differences of 10 per cent correspond to two responses per day. Since, further, the curves are determined by the activity of single animals, the information on the control subjects is significant in revealing how much of the behavior discrepancies may be attributed to individual differences rather than to experimental variables.

In Graphs I through IV (Figure 1) the average incidence of certain responses is plotted as a function of the delay period. The performance of each *NoSC* animal in the standard (*A*) group is indicated by a hollow circle, while hollow squares are used for Group *A* rats trained with the short-circuited contacts. To facilitate comparisons, the averages for the animals in each series have been joined by solid lines. In general, the resulting curves suggest that better conditioning attends shorter intervals of delay. But it is necessary to examine the data in detail before definite conclusions may be drawn.

Leg Withdrawals (Graph I). The general level of conditioning as indicated by this response was quite low. Only 2 of the 16 animals (*A*3 and *A*13) made early withdrawals in as many as 30 per cent of the trials.[2] Although these two were trained at short intervals, 1/3 and 2/3 sec., respectively, too much emphasis cannot be placed upon their records. Six other animals trained at the same delay periods exhibited conditioning comparable to that obtained at longer intervals. These results certainly indicate that the difficulty in establishing a conditioned flexion reflex, as reported in this and earlier papers, cannot be attributed to the use of a short interval of delay.[3]

---

[2] It should be noted that the use of *early* responses as an index of leg responses is especially unfavorable to the short intervals. The shock, coming within 1 sec. after the onset of the buzz, masks any conditioned leg reactions of longer latency.

[3] Whenever identification of an animal is important, his letter and number have been placed on the graphs. Other animals' results are plotted but not numbered, to avoid confusion.

All Leg Movements (Graph II) If movements of active or passive extension are added to those of flexion, and the resulting incidences of "all leg movements" are plotted, the curves are somewhat higher than those representing flexions alone There is a marked rise throughout the long delay periods. It is believed, however, that this rise is in part an artifact. In the holder employed, all animals made rather frequent "spontaneous" postural adjustments in the absence of obvious external stimulation By sheer chance these movements would be more apt to appear during long delays than during short ones Thus, the increased incidence of these minor leg movements does not indicate greater effectiveness of the long periods as the graph would suggest

Sharp Inspirations (Graph III) From the solid-line curves in Graph III, it is clear that, for Group *A* animals, the shorter intervals were most effective for conditioning "sharp inspirations" But because of the observed individual differences between rats it is impossible to determine exactly the optimal interval. With intervals longer than 2/3 sec, conditioned inspirations made by *SC* rats were definitely decreasing The curve for the *NoSC* animals, on the other hand, shows its first major drop at the 4-sec interval There is also some suggestion that the 1/3-sec delay was inferior to 2/3 sec Only one of the 2/3-sec rats failed to surpass all four of the 1/3-sec animals As indicated in Footnote 2, a measure based upon *early* responses is an unfavorable index of conditioning at the short intervals Had unreinforced "test" presentations of the buzz been used, the frequency of "sharp inspirations" at intervals below 2/3 sec might have been greater than that recorded by the present technique.

All Breathing Changes (Graph IV) When one considers all modifications of the rate and form of breathing, the gradient of effectiveness over the different delay periods is no longer apparent This is especially true for the *NoSC* animals of the *A* group These less specific breathing changes are conditioned more readily than the more precise "sharp inspirations," and with this general index the longer intervals appear fully as favorable to learning as do the others. In the opinion of the writers, however, this is not as significant a measure as that involving a sharp inspiration of measured amplitude

SPECIAL PROCEDURES AND THEIR RESULTS (GROUPS *B*, *C*, AND *D*)

The results on the main experimental group indicated that, as measured by the conditioned "sharp inspiration," the shorter delay periods were most effective It seemed desirable to make certain variations in training procedure to test this conclusion under other conditions Twelve more rats were studied in three groups of four each. On the graphs their results are indicated by solid symbols, in contrast to the hollow ones used for the "*A*" animals

Group B The daily records of many of the animals in the standard group indicated a falling off in the amplitude of response as training was continued. This is a rather common occurrence in experiments employing electric shock as a stimulus (10, 13) To counteract this decreased effectiveness, the shock was increased to 3 ma. after trial 10 and again to .4 ma. after trial 15 on the second and third experimental days. During the later days it was adjusted to 4 ma at the outset and maintained at that value throughout the rest of the training A second variation was introduced in the use of four or five isolated (9) or "priming" (5) shocks unaccompanied by the buzz before each session Rats *B*1 and *B*2 were trained with the 2-sec delay, and *B*3 and *B*4 with one of 7 secs As usual, the even-numbered rats were trained with short-circuiting contacts, and the others without them

As measured by leg withdrawals, the performance of the animals trained under these conditions did not differ materially from that of the "*A*" animals. One animal (*B*1) developed a consistent extension of the foreleg as part of a general tonic change. In regard to "sharp inspirations," the 2-sec animals (*B*1 and *B*2) conditioned to the same degree as the corresponding rats of the main group (*A*7 and *A*8) But the animals trained at the 7-sec. delay period (*B*3 and *B*4) showed a much higher incidence of sharp inspirations than did the comparable animals from the standard group (*A*11 and *A*12). Much of this improvement was manifested during the early trials of each session and may be attributed to the "warming-up effect" of the preliminary shocks (10). The significance of these results will be discussed later.

Group C This group was used to study the effects of changing the type of shock employed as the unconditioned stimulus It will be remembered that the stimulating apparatus described earlier supplied a half-wave rectified alternating current supposedly resulting in 60 shocks per sec For the present group, "*C*," a Harvard induction coil was used. The primary was activated directly (without interrupter) from a transformer. The resulting stimulating current presumabily had a fairly sinusoidal form and delivered 120 shocks per sec, free from polarizing effects The induction coil separation was adjusted to produce responses of usual amplitude. Rats *C*1 and *C*2 were trained at the 1-sec delay, while rats *C*3 and *C*4 were trained at 4 sec With the exceptions noted, the animals were trained under the same conditions as were the rats of Group *B*

The results with these rats were surprisingly similar to those obtained from the animals in Group *B* except that *C*2 and *C*4 were very excited and showed frequent leg relaxation At the short interval, *C*1 and *C*2 differed little from the corresponding rats of the standard group (*A*5 and *A*6). But the rats trained at 4 sec. (*C*3 and *C*4) showed a much higher incidence of sharp inspirations than did *A*9 and *A*10 with which they are compared The change in the type of shock certainly had no effect upon the establishment of conditioned leg withdrawals. This result is in agreement with that previously reported by Schlosberg (10), so it is clear that the change to alternating current has no influence on the rate of conditioning For this reason the present data may be grouped with those from the "*B*" animals to furnish further evidence that the procedures ("priming" and increasing shocks) which favor conditioned breathing are more effective on the relatively unsatisfactory long delay intervals than they are on the shorter ones

Group D The fundamental purpose of this particular experiment was to determine whether or not a variable delay period would favor conditioning in the present situation. Each animal in Groups *A*, *B*, and *C* had been trained with a constant, fixed interval between the onsets of buzz and shock Liddell (5) has stated that training with a variable interval favors the appearance of early responses To test this question, two rats (*D*1 and *D*2) were trained with the buzz starting 1/3 to 2 sec. before the shock, and two more (*D*3 and *D*4) with a 1- to 7-sec interval The delay periods were controlled by hand rather than by the timer and were presented in accordance with a prearranged "chance" schedule The average of the short variable periods was 9 sec, and of the long variable 4 sec Rats

*D*1 and *D*2 were trained under shock conditions similar to the "*A*" group, but *D*3 and *D*4 were shocked like the "*B*" rats.

In Graphs I through IV (Figure 1), the results from the "*D*" animals are plotted at the average value of the interval Their level of performance is represented by an arrow or dumb-bell to suggest the variable nature of the interval. Because of the shock conditions the animals of the short variable delay (*D*1 and *D*2) should be compared with rats of the "*A*" group (*A*5 and *A*6) When this is done it is evident that the short variable delay is no more effective than the corresponding fixed interval. The animals at the long variable period (*D*3 and *D*4) are compared with animals of the "*B*" or "*C*" groups Such a comparison shows that the long variable delay animals fall between *B*3 and *B*4 and betweeen *C*3 and *C*4 with regard to early sharp inspirations. One of the animals, *D*4, markedly surpassed all other animals trained with intervals longer than 2/3 sec. when leg flexions are used as the index of conditioning His record might easily be attributed to individual variability, however, since his experimental mate, *D*3, was below the level of comparable animals by the same index There is thus no clear evidence from these data that variable delay periods are more effective than are fixed intervals of similar duration This conclusion should be limited to the rat under the present experimental conditions and may not hold for animals higher in the phylogenetic scale.[1]

GENERAL RESULTS

1. *The Effectiveness of Various Intervals of Delay* The results obtained from the animals of the main experimental (*A*) group indicate that the intervals used are not all equally favorable to the establishment of conditioned sharp inspirations. Such inspirations furnished the most satisfactory index and demonstrated that the 4- and 7-sec delays are clearly less effective for conditioning The exact shape of the gradient cannot be determined from the data available, but there is a suggestion that the decrease in effectiveness is first seen at the 2-sec interval Conditioning (in terms of sharp

---

[1] Further details of these experiments, including protocols for each animal, may be found in the Master's thesis of the junior author, on file in the Brown University Library

inspirations) was also less effective below 2/3 sec. This fall may be apparent rather than real, since some early responses may have been masked by the shock at the 1/3-sec. delay period. The variations in effectiveness as determined for the rats of the *"A"* group are in substantial agreement with those obtained by Schlosberg for the human knee-jerk (8), Wolfle for the hand withdrawal in man (12, 13), and by Yarborough (14) and by Warner (11) for the flight of rats from the ensuing shock. If the conditioned stimulus had been a single stroke of a bell, rather than the continued sounding of the buzzer, the diminished effectiveness of the longer delay periods might be attributed to a waning of the effects of stimulation. Thus Hunter says:

> "... one stimulus (conditioned or unconditioned) must be sufficiently recent to fall within the range of the antecedent and subsequent gradients if it is to be connected with another stimulus (unconditioned or conditioned) by means of learning" (4, pp. 310-311).

But in the present situation the buzz lasts until the shock is presented. Therefore any gradient in effectiveness must occur during the actual duration of the buzz. There is no reason to believe, however, that the buzz is an equally effective stimulus throughout its duration. Indeed, psychologists have long recognized the waxing and waning (cf., *"Anklingen"* and *"Abklingen"*) of the experiences resulting from continuous stimulation. If such waxing and waning of the effectiveness of the buzz may be assumed in the present situation, the more extensive conditioning at the moderately short intervals of delay is not surprising. It is simply another example of the familiar relationship between the strength of the conditioned stimulus and its effectiveness in establishing a conditioned response.

2. *Effect of Preliminary and Increasing Shocks.* When conditions were modified by the use of preliminary shocks and shocks of increasing intensity, the gradient in conditioned sharp inspirations disappeared. The modified technique was much more effective than the standard one at the 4- and 7-sec. intervals, but yielded about the same level of conditioned inspirations as the standard method at the 1- and 2-sec. delay periods. This might be understood by considering the level established by the short delay periods to be near a "Physiological Limit" in regard to respiratory changes. Thus

there would be more chance for the modified conditions to show their effectiveness at the long intervals where conditioning was clearly below this limit But, if we consider other indices of conditioning, as the leg withdrawal, it is clear that the animals of this experiment were well below the limits of conditioning described elsewhere (10). The changed shock conditions which increased the incidence of sharp inspiration did not modify the level of leg reactions

A more promising approach may be made if we consider that the *B*, *C*, and similarly trained *D* animals were generally more excited, as indicated by intermittent squealing and struggling between trials, than were the animals of the standard group. The excitement may very well have delayed the subsidence of the central effects of stimulation so that they were as strong near the end of the buzz as they had been at the beginning This statement does not imply any special view of cortical functioning, for continued central activity may be the result of circular reinforcement from the responses of muscles and glands. Indeed, continued widespread response to stimulation constitutes excitement as it is used here. If the effects of stimulation by the buzz last longer under the changed conditions, it is easy to see why the long intervals would be as effective as the shorter ones The low record of *B2* does not fit this view. If 10 rats instead of 1 were represented by each point on the graphs, it might be possible through graphical analysis to infer a second process which became effective under the changed conditions for the 4- and 7-sec. delay intervals. But in view of the marked variability of the rats it seems preferable to avoid such an interpretation, and to adopt one which is less dependent upon single animals. The view that has been suggested above is reasonably consistent with the results obtained from 9 of the 10 rats trained under the modified procedure (*B*1, *B*3, *B*4, *C*1, *C*2, *C*3, *C*4, *D*3, *D*4). It may, therefore, be accepted in a very tentative fashion subject to later verification.

3. *Effects of the Short-Circuiting Contacts* In the first major paper of this series (9), Schlosberg reported that the inclusion of contacts so arranged that the shock was terminated or prevented by the tail withdrawal did not increase the degree of conditioning. Similar results with the leg withdrawal were presented in the second publication (10). In each paper, mention was made of the fact that this finding was not what one might expect from the "law of

effect " The animals did not learn more rapidly in the situation in which "success" in avoiding the shock was possible It was thought desirable to obtain further data on this question with respect to varying periods of delay and so the animals, with the exception of the four controls, were all trained in pairs. The odd-numbered animals received a shock of fixed duration (1/3 sec), while the even-numbered animals terminated each shock if they withdrew their legs and maintained the flexion until the termination of the buzz.

A review of the graph which presents the data on leg withdrawals or flexion indicates that six of the *SC* animals made these responses more frequently and six made them less frequently than their *NoSC* mates The average incidence of this reaction for the *NoSC* group was 14 per cent and for the *SC* group 13 per cent But, in these withdrawals, the *SC* animals rarely avoided the shock, even though it was theoretically possible An examination of the kymograph records showed that no animal made the sustained contraction necessary for avoidance of the shock in more than three of the 125-150 trials In some cases there were more than this number of accidental omissions of the shock Since the "successful" withdrawal was so infrequent, the law of effect is inapplicable to the behavior of the animals reported in this study

When *SC* and *NoSC* animals are compared on the basis of sharp inspirations, the *NoSC* technique seems to be more effective. In the 12 pairs, only four of the *SC* animals surpassed their mates and this was by a small margin in each case The average incidence of sharp inspirations in the 12 *NoSC* animals was 78 per cent, that for the *SC* group, 59 per cent It is clear then that the use of the *SC* contacts makes for a less consistent conditioned sharp inspiration. This result might have been expected since the leg reaction in this situation shortened the shock duration to about .1 sec. The *SC* animals received a shock which lasted only a third as long as that delivered to the *NoSC* rats and in consequence did not condition as readily The differences in this study are more marked than those reported previously by Schlosberg, largely because the shock of fixed duration was 1/3 sec in the present studies as contrasted with 1/6 sec. in the earlier studies. In the behavior of the animals, the difference between 1/3 and 1/10 shows up more clearly than that between 1/6 and 1/10. The longer shock, through summation of its elements, has a stronger effect on the animal and brings more extensive conditioning.

4 *The Nature and Development of the Responses.* In the foregoing sections most attention has been directed to leg withdrawals and sharp inspirations, the two responses that furnished the best indices of conditioning In this section, all of the responses will be considered in an attempt to form some impression of their nature.

The first feature of the response to appear was usually the breathing response. It began as a minor irregularity in the respiratory pattern and was soon displaced by the sharp inspiration, usually accompanied by a squeal With some of the animals, this squeal eventually became the first of a series of squeals which continued throughout the entire period of delay While the breathing reaction was developing, many of the rats also developed a conditioned tail response, generally consisting of a stiffening rather than a flailing of this member Leg reactions were the last to appear and often consisted largely of relaxations or extensions rather than of specific flexions The withdrawals which did occur were most frequent in the early sessions, occurring as part of a struggle or startle response. It may be seen that these responses are all diffuse, rather than specific, and that the least precise ones developed first. The opposite of this order was observed during extinction series. Seven animals (*A*3, 4, 9, 10, 11, 12, 13) were subjected to a series of unreinforced buzzes at the conclusion of the sixth training session. The conditioned stimuli were of their usual duration for each animal and were separated by 30-sec. intervals One rat (*A*3) showed no decrement whatsoever after 25 stimulations. The other six showed a progressive weakening of the breathing response and a falling-out of the skeletal reactions until the conditioned response was almost completely eliminated by the 25th or 30th unreinforced presentation. Two of the animals gave more responses during the extinction series than they had during the regular training. This increased activity, however, was short-lived and did not refute the general statement that the least specific reaction, the first to appear in learning, was the last to be extinguished

The most specific of the responses to shock, the leg withdrawal, did not become conditioned to any appreciable extent in the six experimental sessions. In general, the animals' behavior was typical of the early stages of the conditioned reactions as observed in higher animals (as dogs) Certain of the responses were intimately related to excitement Of the 28 animals employed, the 7 most excited, as judged by the frequency of squealing and struggling between trials, were among the 10 animals which made most frequent tail reactions Similarly, six of these seven excited animals were included in the group which demonstrated "regular squealing" during the sounding of the buzzer As noted above, much of the improvement attending the use of preliminary and increasing shocks may be attributed to the greater excitement induced by these conditions

Additional evidence of the non-specific nature of the conditioning and its relation to excitement is seen in marked generalization on the stimulus side Tactual stimulation on the back or legs of the animal invariably released typical inspirations, squeals, and struggling in which the leg and tail participated Weaker stimuli, such as tapping on the table, often elicited squealing

Still another indication of the unstable level of conditioning was the rare development of a patterned reaction, that is, a response in which any one breathing change was consistently accompanied by the same leg and tail reactions from trial to trial Only three animals (*A*3, *B*1, and *D*3) exhibited such patterned behavior

Finally, although the temporal features of the situation made it one of the "delayed reflex" type, genuine delayed responses were almost never observed The conditioned reaction typically began immediately after the onset of the buzz and was not related in its time of appearance to the actual delay period

These results, all showing that conditioning never developed beyond a relatively early stage, are quite in keeping with those previously reported by Schlosberg (10) Conditioning progresses somewhat further in these rats than Liddell reports for rabbits trained under comparable conditions, but not nearly so far as in Liddell's sheep, pigs, or goats (5) In experiments where rats are allowed to run from shock (1, 3, 11) much more rapid and specific learning occurs The difference would seem to be in the combination of shock and restraint. These conditions produce in the rat a state of excitement which is not conducive to the development of precise responses. Schlosberg has suggested a similar interpretation in an earlier paper (10), and the results of the present experiment support this view in every way.

Considering the series of papers as a whole, it may be seen that many experimental factors have been varied but no set of conditions has yet been found under which a specific withdrawal of the rat's foreleg may be consistently elicited by a new stimulus (light or buzz) associated during training with the shock. It now seems clear that the general situation is unfavorable for such conditioning with the white rat.

SUMMARY AND CONCLUSIONS

1. Twenty-eight rats were trained in a simple "delayed reflex" type of conditioning situation in which a buzz was followed by an electric shock to the right foreleg. Six sessions of 20 trials each comprised the training period.

2. Each animal was trained at one of eight periods of delay (1/3 sec., 2/3 sec., 1 sec., 2 sec., 4 sec., 7 sec., short variable, and long variable).

3. Under "standard" conditions, the 2/3- and 1-sec. intervals were more favorable to the development of conditioned sharp inspirations than were the longer delay periods. This result is attributed to a waning in effectiveness of the buzz as it continued over a period of several seconds.

4. Under modified ("stronger") shock conditions, there was no clear gradient of the effectiveness for the several intervals, as measured by the same index. Such shock conditions led to excitement which apparently delayed the waning in effectiveness of the buzz.

5. The frequency of conditioned leg flexions was very low under all conditions. Variable delay periods were no more effective than were fixed ones in eliciting this reaction.

6. Due to the infrequency of withdrawals which resulted in the complete avoidance of the shock, the law of effect is inapplicable to the leg withdrawal reported in this paper. But the early termination of the shock by leg withdrawal diminished its effectiveness in establishing conditioned breathing reactions.

7. Breathing, tail, and leg reaction developed in the order named, and fell out in reverse order during extinction.

8. The results of this and of earlier papers indicate that the white rat, when restrained and shocked, can readily develop diffuse responses More specific reactions, as leg withdrawals, are rarely acquired

References

1 CARR, H A & FREEMAN, A S Time relationships in the formation of associations *Psychol. Rev*, 1919, **26**, 465-473.

2 HILGARD, E. R Conditioned eyelid reactions to a light based on the reflex wink to sound *Psychol Monog*, 1931, **41**, No 184 Pp 50

3 HUNTER, W S Basic phenomena in learning *J Gen. Psychol*, 1933, **8**, 299-317

4 ——— Conditioning and extinction in the rat *Brit J Psychol*, 1935, **26**, 135-148.

5 LIDDELL, H S, JAMES, W T, & ANDERSON, O. D The comparative physiology of the conditioned motor reflex. *Comp Psychol Monog*, 1934, **11**, No. 1 Pp 89

6 PAVLOV, I P Conditioned reflexes. (Trans by G V Anrep) London Oxford Univ. Press, 1927 Pp xv+430

7 RAZRAN, G. H S Conditioned withdrawal responses with shock as the conditioning stimulus in adult human subjects *Psychol Bull*, 1934, **31**, 111-143

8 SCHLOSBERG, H A study of the conditioned patellar reflex. *J Exper Psychol*, 1928, **11**, 468-494

9 ——— Conditioned responses in the white rat *J. Genet. Psychol*, 1934, **45**, 303-335

10 ——— Conditioned responses in the white rat II Conditioned responses based upon a shock to the foreleg *J. Genet Psychol*, 1936, **49**, 107-138.

11 WARNER, L H. The association span of the white rat *J Genet. Psychol*, 1932, **41**, 57-90

12 WOLFLE, H M Time factors in conditioning finger withdrawal *J Gen Psychol*, 1930, **4**, 372-378.

13 ——— Conditioning as a function of the interval between the conditioned and the original stimulus *J Gen. Psychol*, 1932, **7**, 80-103

14 YARBOROUGH, J. U The influence of the time interval upon the rate of learning in the white rat *Psychol Monog*, 1921, **30**, No 135 Pp 52

University of Rochester
Rochester, New York

Brown University
Providence, Rhode Island

REACTION-TIME APPARATUS*

Brown University

HAROLD SCHLOSBERG

In recent years a rather large number of new chronoscopes have been developed. Those who developed the new instruments have usually presented some data on their accuracy and constancy, but relatively little information is available in regard to the way these instruments actually perform in the hands of other experimenters. Dunlap (3) has reported such data on a number of chronometric devices, but he did not test the recent stop-clocks (1), the vacuum tube chronometer (4) or the older and excellent Delabarre Synchronous Motor Chronograph (2) Further, his results on his own instrument were obtained from earlier models, rather than from the more recent Stoelting instruments Dallenbach's (1) report on the stop-clocks does not present data in regard to the absolute error of these convenient instruments, but merely indicates their variability To fill in these gaps, the present writer tested nine instruments of seven different types, using constant intervals, and comparable conditions for each instrument. The results of these tests are presented in the present paper

To furnish the constant intervals, a simple gravity timer was constructed To adapt this timer to certain of the instruments it was necessary to devise a balanced relay Both timer and relay will be described before the tests on the chronoscopes are presented. In passing, it should be noted that the balanced relay makes it possible to use the wall model of the stop-clock as a demonstration instrument, in conjunction with any or all of the Dunlap equipment

A SIMPLE GRAVITY TIMER

No attempt will be made here to discuss in detail the advantages and disadvantages of various devices for obtaining fixed intervals of time The arrangement to be described has the advantage of simplicity of construction, and is reasonably constant It consists

---

*Received in the Editorial Office on April 24, 1936

essentially of a weight which falls along a loosely-fitted brass rod and briefly interrupts several contacts as it passes them. Details of its construction may be obtained from Figure 1. This figure also

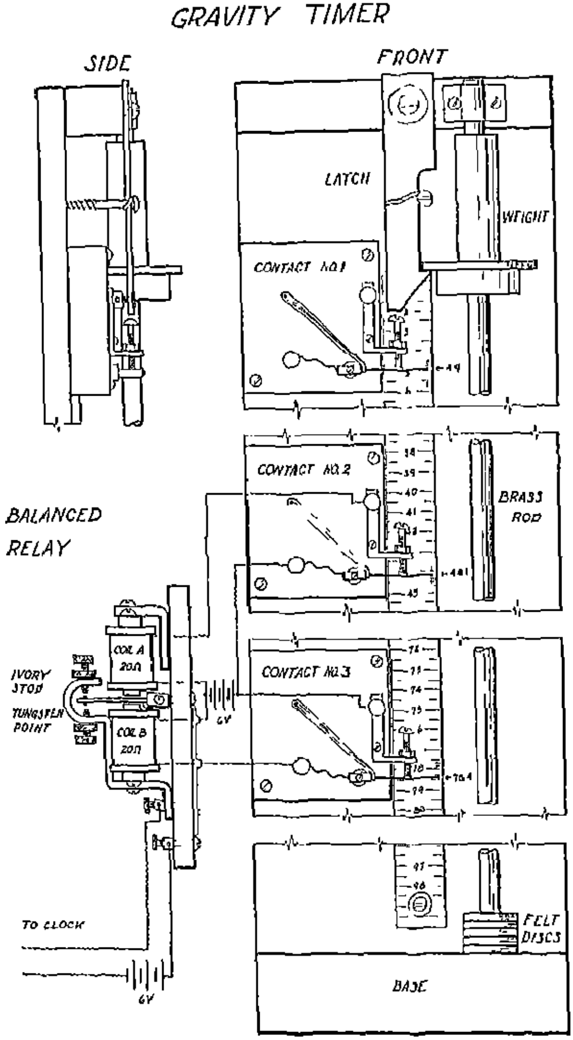


FIGURE 1

shows the balanced relay and connections appropriate for the control of the stop-clock. The meter stick shown on the timer may be used as a scale for determining the dimensions of the various parts. A curve (based on $g = 980$ cm / sec ^{2}) was constructed, and the contacts were placed at the following distances and (theoretical) temporal intervals from the starting-point: No. 1 at 4.9 cm. or 1 sec., No. 2 at 44.1 cm. or 3 sec., No. 3 at 78.4 cm. or .4 sec. With this arrangement the weight already had considerable velocity before it hit contact No. 1, so that a variation of 1 mm. in the placement of this contact would introduce an error of only 1 ms. A similar displacement of contacts 2 or 3 would cause errors of 1/3 ms. or of 1/4 ms., respectively. Contacts could be placed within 1 mm. of the calculated position, and did not vary with use. Additional errors were introduced by the resistances of the contacts, and by friction between the hard-rubber bushing of the weight and the polished brass rod. It was necessary to determine the magnitude of these errors before the instrument could be used. This was done with the Delabarre Synchronous Motor Chronograph (2), which will be described below. The interval between contacts 1 and 2 was found to be 308.7 ms., instead of 300 ms., and the shorter interval between contacts 2 and 3 measured 102.05 ms., instead of 100 ms. The variability of the timer could not be determined by the chronograph. Results with the Dunlap Chronoscope (to be presented below) indicated that the mean variation of the shorter interval was less than 5 ms., and of the longer interval, not more than 1 ms. In summary, the short interval may be taken as 102 ± 5 ms., and the longer one as 308.7 ± 1 ms.

THE BALANCED RELAY

As indicated in the Introduction, this relay was built to convert the break-break of two different circuits, as used with the Dunlap Reaction Board, into a closing and opening of the circuit through the clutch of the stop-clock. In practice it was found very useful in testing other instruments which employ a make-break of the controlling circuit. A combination of the relay and the gravity timer described above made it possible to test every time-measuring instrument in the laboratory.

The relay and its connections are shown in Figure 1. It is simply an inexpensive relay (Novel Form Pony Relay, J. H. Bunnell & Co., N. Y.) to which has been added a second pair of coils, in place of the spring. To insure an adequate match, the additional pair of coils was taken from another relay of similar design. When the circuit through magnet *A* is briefly interrupted, the armature moves to one side, closing the secondary circuit. This movement introduces a latency of 7-9 ms. When the circuit through coil *B* is broken, the armature returns, opening the contacts with no apparent latency. Thus, the relay decreases the actual interval by approximately 8 ms., depending slightly on the distance the armature is free to move. These characteristics were measured by recording the changes in the various circuits, together with a 120-cycle time line[1] on a rapidly moving kymograph drum. The relay has worked consistently for several thousand readings, and has been used with Dunlap Voice Keys, as well as with the timer. When the relay is to be used with the Dunlap Reaction Board, the two magnets should be connected in place of the chronoscope magnets. If it is desirable to keep the Dunlap Chronoscope in the circuit, chronoscope and relay magnets may be connected in parallel. Since there are other resistances in the circuits, the resistances of the relay coils should be high enough so that they will not shunt too much current from the chronoscope coils. The magnets of the Dunlap Chronoscope described below were found to have a resistance of about 7 ohms. With 20 ohm coils in the relay, the Dunlap Chronoscope worked in a normal fashion, if the battery was fully charged. In passing, it should be pointed out that in the tests to be described below, in which the gravity timer was employed, the resistances of the various instruments were not particularly significant, for they were all large in comparison to those of the timing contacts and leads.

The Delabarre Synchronous Motor Chronograph

This instrument was described by Professor Delabarre (2) 14 years before Dunlap first described his instrument. It employs a kymograph drum, driven by a synchronous motor at the rate of one revolution in two seconds. The length of the interval to be measured is determined by measuring the length of a line drawn on the smoked surface by a marker. For simple auditory reactions the stimulus is furnished by the click of a wooden point as it falls on the writing surface. The response is the depression of a lever which directly lifts the writing point from the paper. Assuming the alternations of the controlled A C to be accurate,[2] there would seem to be no possibility of a constant error in this measurement. For recording other types of stimuli and responses, a light electric marker is used. The armature of this marker is mounted so that breaking the circuit through the coils permits the point to leave the paper, to which it returns after the circuit is reestablished. The break in the recorded line is much clearer than the more usual lateral deflection, for the tension of the writing point on the smoked surface gives it an initial acceleration, instead of retarding its movement.

---

[1] The time line was produced by the tuned vibrator described by the present writer (5). It is no larger than a standard marker. Compare Dunlap's comment on a similar instrument (3, p. 7).

TABLE 1

SUMMARY OF THE DATA OBTAINED FROM FIVE CHRONOMETRIC DEVICES

Each entry is computed from a group of 10 successive readings

Trials	Delabarre Chronograph Av	Delabarre Chronograph M V	Hipp Chronoscope Av	Hipp Chronoscope M V	Dunlap Chronoscope Av	Dunlap Chronoscope M V	Pendulom Chronoscope* Av.	Pendulom Chronoscope* M V	S E. T. Co Stop-clock* Av	S E. T. Co Stop-clock* M V
		A	At the short interval (102 ms)							
1-10	102 3	1 20	102 1	1 54	99 1	1 12	126 0	2 90	103 2	1 8
11-20	102 6	88	101 8	88	97 4	72	122 9	2 50	107 1	3 52
21-30	101,8	48	101 3	76	96 4	56	122 1	1 52	111 0	3 0
31-40	101 5	,80	100 8	1,04	97 3	70	123 6	1 00	108 3	3 62
Average	102 1	84	101 5	1 06	97 5	78	123 7	1 48	107,4	2 99
		B	At the long interval (309 ms)							
1-10	309 3	2 10	303 6	2 08	302 6	88	309 7	1 30	323 1	4,88
11-20	309 2	1 80	302 6	2 48	301 5	1 30	307 1	1 48	318 0	3 60
21-30	308 5	1 50	304 1	1 72	302 4	1,36	309 6	1 68	314 9	2 48
31-40	307 6	2 32	299,7	1,57	300 5	70	304 7	2 36	312 3	3 96
41-50			304 0	1 60	301 8	1 12	310 5	2 60	315 6	2 88
Average	308 7	1 93	302 8	1 89	301 8	1 07	308 3	1 88	316 8	3 56

*These readings would have been 8 ms greater if the relay had not been used.

---

[2] A member of the Engineering Department of Brown University estimates their accuracy to be better than 1 per cent.

For measuring the intervals produced by the gravity timer, all three pairs of contacts were connected in series through the marker and a single storage cell As the weight passed each contact the writing point left the smoked surface for a brief time (20-40 ms), so that a single record would show a line with three breaks in it Forty such records were taken, and the distances between breaks 1-3 and 2-3 were measured. The average of these measurements (Table 1) showed the long interval (contacts 1-3) to be 308.7 ms in duration A similar average for the short interval (contacts 2-3) was 102 05 ms Since the drum was turning at constant speed, and since measurements were made from break to break of the same marker, there would seem to be no possibility of a constant error in these averages. Further, the results are clearly what might be expected, for they show that friction between the falling weight and its guide rod decreased the speed of fall by approximately 2 per cent. Curiously, a change of this magnitude had been predicted by a trained physicist who had seen the gravity timer before it was tested

The variability of the readings was more pronounced The mean variation for the shorter interval was .84 ms, and for the longer one, 1.9 ms. These errors are less than 1 per cent of the intervals measured, and are due largely to errors of interpolation in reading the records. The measurements were made on a scale on which 10 ms = 1/8 inch. Such variability is not a serious disadvantage for most psychological studies, and is more than counterbalanced by the absence of a constant error. In fact, this chronograph is the only one of those tested that the writer would trust without frequent calibration, if any precision was desired. The slight inconvenience involved in preparing the drum can undoubtedly be decreased by the use of different recording surfaces, and the trouble of measuring the records is scarcely greater than that involved in subtracting successive readings of the Hipp One of the major advantages of the chronograph lies in certain ingenious arrangements which make it possible for a person to measure his own auditory reaction-times without the aid of another person. In short, this rather neglected instrument has many desirable features

The Standard Electric Time Company's *SW*61

This instrument was tested simultaneously with three others (Dunlap, Hipp, and Pendulum) so that the results, as presented in Table 1, are strictly comparable for all four instruments. The remaining instruments used in this study were tested separately, since it was not possible to wire them all into the same circuit.

The synchronous motor of the wall model of the electric stop-clock was connected to the regulated lighting circuit. The clutch was connected to a well-charged 6-volt storage battery, through the contacts of the balanced relay, as shown in Figure 1. Forty readings were taken at the short interval, and 50 at the longer one. The dial of the clock is divided into .01-sec. intervals, but the position of the hand was estimated and recorded in .001 sec. The averages and mean variations of successive groups of 10 readings are presented in Table 1. For the short interval (102 ms.), the average of the means was 107.4 ms., and of the mean variations, 2.99 ms. The range was 98-118 ms. For the long interval (309 ms.) the average of the means was 316.8 ms., and of the mean variations, 3.56 ms., while the range was 304-320 ms. At least 2.5 ms. of the variability must be attributed to the clock. This value compares favorably with one of 4 ms. reported for a similar instrument by Dallenbach (1). The constant error of 8 ms. for the long interval and of 5 ms. for the short one represents the over-all efficiency of clock and relay. As noted above, the relay had a latency of 8 ms., so if the clock had been used without the relay its error would have been approximately +15 ms. This constant error is due to the clutch mechanism in which an electromagnet is opposed by a spring. Further, the error might be expected to be different for different clocks, or for the same clock under different conditions of current, temperature, and age. The error is a terminal one, and would not increase with the length of the interval measured. It can be eliminated by removing the back and tightening the spring which opposes the magnet, but perhaps a simple substraction of 15 ms. (or 7 ms. when the relay was used) would be more convenient.

The stop-clock is an extremely convenient instrument. It may be reset by pressing a remote-control button. As Dallenbach has indicated, the hand does not always return to zero, but with the particular instrument reported here the reset is always complete if the button is depressed for a full second. The motor may run at half speed for a few minutes on a cold day, but it soon warms up and runs at correct speed This may be due to the fact that the Telechron motor is mounted on its side, causing increased friction The clock is constant enough for ordinary class work, and is particularly valuable for large demonstrations, because of its 10-inch dial Demonstrations of reaction-time measurement have been difficult in the past, for they necessitated a specially-built instrument, such as Wundt's (6) With the stop-clock and balanced relay such demonstrations are simple and inexpensive.[3]

A small model (*S*1) with A.C. clutch and a dial reading to 1/300 sec was tried out It was quite variable and hard to read The later *SC*1 described by Dallenbach (1) seems to be more constant The latest model has a more convenient dial, reading to .01 sec., and is probably more constant than the wall model.

THE DUNLAP CHRONOSCOPE

The instrument tested was one of Stoelting's *Type C*, purchased in 1927, and modified at a later date to read in ms when used with rectified A.C It has had fairly heavy intermittent use since its purchase, and has been overhauled by the Stoelting Co on one occasion. When it was repaired, a felt washer was added to the shaft, in an attempt to keep oil out of the clutch plate

The average of the first 10 readings at the short interval was 97 ± 52 ms. The third group of readings gave an average of 90.7 ± 2.36 ms. Previous experience had indicated that this type of change was caused by oil on the clutch So, despite the fact that the clutch had been cleaned on the previous day, the chronoscope was taken completely apart and cleaned Dunlap (3) has suggested the occasional insertion of a piece of paper between the plates, but this procedure is difficult with this particular model, and does not give very lasting improvement The chronoscope was then put back into operation, and the results included in Table 1 were obtained They represent the performance of this particular instrument under optimal conditions. For the short interval the mean of the four averages was 97 52 ms ± 78 ms. For the long interval a corresponding average was 301 8 ms ± 1 07 ms. Most of the variability may be attributed to the timer But there seemed to be a constant error of approximately 5 ms In view of Dunlap's statement (3) that "properly adjusted and connected, the instrument shows no error," it seemed desirable to find the source of error. It could hardly be attributed to the gravity timer, for friction or contact resistance cannot make the weight fall *faster* than 980 cm /sec.^{2} The three sources of error that Dunlap suggested in the same paper were also checked, with the following results

---

[3] Very effective voice keys, similar to Dunlap's, may be made rather easily if cellophane is used for the membranes A pair of extra voice keys makes it possible to use the clock as an additional instrument when an experimental group is studying association-time. As a matter of fact, voice keys are probably no more accurate than lip keys, or even hand keys, for such purposes

1 *Unbalanced Circuits* Any lack of balance must have been in the instrument itself, for the resistance of contacts and wiring were insignificant The resistances of the coils of the instrument were tested and found to be essentially equal

2 *Oil on the Clutch* The whole motor had just been cleaned, and no additional oil was added to the front bearing. Hence, no oil could reach the clutch This difficulty is a serious one, and makes the chronoscope unreliable if more than an annual drop of oil is used on the front bearing

3 *Wear on the Front Bearing* Dunlap's test was hard to apply on this instrument, because of the additional gearing But this trouble seems the most likely cause of error, for the variability of the instrument was much greater if the hand was not reset after each reading Apparently the initial point from which the hand starts has an effect on the operation of the clock

4 *An Additional Source of Error.* Not mentioned by Dunlap, it may be introduced by the play in the additional gearing made necessary by the change in the dial noted above The hand could be moved through 4 ms when the forward magnet was holding the clutch plate on the brake For all the readings presented in this paper, the hand had been reset by a backward movement (decreasing readings) for it always stopped with the play taken up in this direction If this precaution had not been taken, the variability would have been much greater

But, whatever the cause of the errors, it is clear that this particular instrument was giving values some 5 ms. shorter than the actual durations of the test intervals, even when operated with great care. As the instrument is apt to be used in a laboratory for routine measurement, it yielded very inaccurate results. This statement should not be interpreted as a criticism of the principle of the Dunlap Chronoscope, but only of one sample of this particular model. Further, when the instrument is used with the reaction board, the error may be much greater than when tested alone. The interval between the presentation of the stimulus and the starting of the chronoscope depends upon the speed with which the compound key is depressed.

THE HIPP CHRONOSCOPE

The record of this instrument (Peyer, Favarger "Newer form") will be presented largely for comparative purposes. Before it was tested at the short intervals it was arranged for operation with balanced magnets, and connected in parallel with the Dunlap. Several runs of 60 sec. each showed that the Hipp was running between 1 per cent and 2 per cent slower than the Dunlap. At long intervals such as these, small terminal errors are of no consequence, so it was clear that the reed of the Hipp was vibrating too slowly. An attempt was made to correct this error by adjusting the damper on the reed, but it met with indifferent success. It is probably more simple to make a numerical correction for this error, as Titchener has suggested (6). The instrument was then tested on the two fixed intervals. By referring to the results (Table 1), it may be seen that the short interval gave an average reading of 101.5 ± 1.06 ms., and that similar figures for the long period were 302.8 ± 1.89 ms. These figures suggest a slight terminal error of 1 or 2 ms., in addition to the running error of 1 or 2 per cent. If correction is made for the running error, the Hipp is probably slightly more accurate, but also more variable than the Dunlap. Much of the variability is to be attributed to the escapement, for similar variability was noted during the 60-sec. tests. Furthermore, in spite of great care in starting the motor, it will occasionally run at half speed without any change in its sound. One such reading was obtained in the series recorded, but has been omitted from the averages. But, in spite of these disadvantages, the instrument is far from obsolete. If the present writer were buying a new chronoscope for its utility, he would probably not choose a Hipp. But, on the other hand, he would hesitate to discard one he already had in place of a more recent instrument.

THE PENDULUM CHRONOSCOPE

The first instrument was purchased from Willyoung (Philadelphia) many years ago. It is calibrated in 01 sec. From the results presented in Table 1 it may be seen that its variability is about the same as that of the Hipp. The absolute value of the readings is essentially correct at the long interval, but 25 ms. too high at the short one. The speed of the pendulum may be changed by moving the counter-weight, but such modification would change only the point at which the readings were correct, unless the scale could be moved. This cannot be done. The instrument can probably be adjusted so that a correction factor in sigma can be determined. For routine work the pendulum chronoscope is very convenient. Its chief disadvantage lies in the coarseness of the scale, particularly at the ends.

A second model of this instrument (Leeds) makes one swing in 2 sec., instead of in one. Since the intervals available on the timer were too short for this range, the pendulum chronoscope was attached to the Dunlap, through the balanced relay, and tested. The readings will not be reproduced, for they cannot be treated in a manner comparable with that used for the others. For intervals of 1600-2000 ms. the pendulum was in substantial agreement with the Dunlap, but read 20 ms. too high around 300 ms. The comments that were made in the preceding paragraph may be applied to this model.

THE CENCO IMPULSE COUNTER

Two early models of this instrument were purchased in 1931. They have been returned to the factory for repair at fairly frequent intervals, for they get out of adjustment rather rapidly when used at 60 cycles (120 impulses per second). Neither instrument had been used since the last trip to the factory. Each one was connected to the timer through the balanced relay, and tested at various voltages supplied by a toy transformer. The first instrument gave 10 successive readings of 36/120 sec. for the 300 7-ms interval, when operated at 16 volts At 18 volts 2 of 10 readings were 1/120 sec too long. At 21 volts the instrument skipped badly, and failed completely at 24 volts. The counter was then tested on the other side of the optimal voltage. Out of 10 readings at 14 volts, one was a scale division too long, and another a unit too short. Similar results were obtained at 12 volts, but below this value the instrument again skipped and stopped On the 97-ms interval, at 16 volts the readings alternated between 11/120 and 12/120, with an average of 97 ms Thus, at its best voltage this counter was accurate beyond one scale unit

The second of these two instruments was more variable At 14, 16, and 18 volts the readings for the 300-ms interval varied between 35 and 37 units, but the average of 10 readings was within 1 or 2 ms of the true value of the interval At 21 volts the average jumped to 340 ms., and at 23 volts the counter started to skip badly At 12 volts the average of the readings dropped to 290 ms, and at 10 volts the instrument stopped completely The average of 10 readings at 14 volts for the 94-ms. interval was correct, but the individual readings varied between 11 and 13 scale units.

These results are in agreement with those of Dunlap (3) in indicating that there is variation from instrument to instrument. If the counters are in good condition, and used at the optimum voltage for each individual instrument, they are satisfactory for rough reaction-time studies. But they become very erratic after moderate use at 120 impulses per sec., and must be returned to the factory. The manufacturers claim that this is not true of the later models. But it would seem advisable to check these counters frequently whenever they are used for reaction studies

VACUUM TUBE CHRONOMETER

The instrument tested is the original one described by Jasper and Andrews (4) The successive opening of two circuits results in a deflection of the needle of a zero-center microammeter By reference to the published calibration curves the deflections can be converted to temporal units By turning a switch any one of five different sensitivities is available. At the greatest sensitivity the 50-unit half-scale of the meter is covered in 5 ms, while at the lowest sensitivity the same range corresponds to 5000 ms. To test the chronometer with the gravity timer, it was necessary to modify the contacts so that they would stay open after the weight had passed. Therefore, small strips of celluloid were arranged so that they would slip between the contact points as soon as they opened and hold them separated.

Twenty readings were taken at the 102-ms. interval, with the chronometer set at sensitivity 3 (range = 105 ms.). Ten readings taken with the needle swinging to the right of the zero point averaged 48.4 scale units, and a similar number to the left averaged 49.2 units. The average of these 20 readings was 48.8 units, which corresponded to 100 ms., indicating an error of —2 ms. It should be noted that the error is only one-half a scale division. (The 50-unit scale has 25 divisions, each one indicating 2 microamps.) Only one of the 20 readings was as much as one scale division from the average.

The instrument was then set to sensitivity 4 (range = 500 ms.). At the 102-ms. interval the average of 20 readings was 13.4, corresponding to 96 ms., with an error of —6 per cent. No reading varied from this average by more than one-half a scale division. At the 309-ms. interval the average of 20 readings was 35.85 scale units, or 295 ms., showing an error of —5 per cent. Eighteen of the 20 readings were 36 units. The contacts were then reversed, and six readings taken at each of the intervals. Under these conditions different calibration curves are used. The errors were almost identical with those noted above.

In summary, the instrument showed an error of —2 per cent at sensitivity 3, and one of —5 per cent at sensitivity 4. It had not been calibrated for the year preceding these tests, and had been used extensively during this period. Further, one resistor had been replaced without calibration. The errors are partly a result of the aging of the tubes. The poorer performance at the lower sensitivity is due largely to the use of carbon resistors. The instrument was originally built to measure intervals up to 50 ms., so wire-wound resistors were used for only the shorter ranges. At these short ranges the instrument is undoubtedly far superior to any of the others described here, for its errors are only a small percentage of the interval measured, rather than absolute ones. If good condensers and resistances were used in all ranges, the accuracy and constancy would depend almost entirely on the care with which the meter was read. Finally, the chronometer is extremely flexible, convenient, and absolutely silent It is the only one of the instruments described that could be used for visual or tactual reactions without some sort of soundproofing.

SUMMARY AND CONCLUSIONS

1. With the addition of a balanced relay, it is possible to use a wall-model stop-clock in conjunction with any of the Dunlap equipment The stop-clock is particularly useful for classroom demonstrations. Its cost and convenience make it a very desirable instrument for such purposes.

2 A simple and constant gravity timer for calibrating chronoscopes was built and used for testing nine instruments of seven different types The results presented in this paper should be considered as measurements of individual instruments, and should be generalized with great caution

3. The only instrument without an error was the Delabarre Synchronous Motor Chronograph

4. The Dunlap Chronoscope had the lowest variability of the instruments tested, but had a constant error of —5 ms. The model tested has the possibility of much more serious errors

5. For careful work all the remaining instruments should be checked frequently on intervals of known duration When properly controlled, each of these instruments has its own advantages and limitations.

REFERENCES

1. DALLENBACH, K M Two new A C. chronoscopes. *Amer J Psychol.*, 1936, **48**, 145-151

2. DELABARRE, E. B. Description of new apparatus. *Psychol Rev*, 1903, **10**, 171-172

3. DUNLAP, K. Chronometric devices in psychological research *J Gen. Psychol*, 1936, **14**, 3-30.

4 JASPER, H. H, & ANDREWS, H L A multi-range vacuum tube chronoscope. *J Gen Psychol*, 1936, **14**, 248-256

5 SCHLOSBERG, H A time marker for 120 cycles *J Gen Psychol*, 1932, **7**, 483-485

6 TITCHENER, E B Experimental psychology Instructor's manual, quantitative. New York Macmillan, 1923 Pp clxxi+453

Brown University
Providence, Rhode Island

ANGER ADJUSTMENTS IN RELATION TO INTELLIGENCE AND ACHIEVEMENT*

Psychological Service Center, St Louis

H Meltzer

When we get angry we betray the fact that we are inferior [1] This is a typical Adlerian statement made by Lazell in his *The Anatomy of Emotion* (6, p 105) Anger, he tells us, "afflicts only those who refuse to permit themselves to become conscious of their inferiorities or failures" This statement implies a positive relationship of anger with inferiority, or with a sense of inferiority, or with both.

In his *Thinking Machine*, Dr Herrick (5, p 212), a neurologist, in discussing anger, talks of the "real mastery of passion by intelligence and will power, the domination of the thalamic by the cortical brain." This implies a negative relationship between anger and intelligence, as well as between anger and integrity of personality

Barbusse (1, p. 83), in the earlier part of his book on Stalin, when building up an argument justifying the cruel practices of Stalin and his followers in dealing with the Opposition, advocates the use of anger as a means of obtaining and maintaining social justice "Actually and obviously," says he, "it is one's duty to strike down a fellow-creature to save a thousand, to save a hundred thousand, to save the future, and to build a better world in which man will no longer be the slave or the victim of man" However, nearer to the end of the book, when he considers his case built and socialism safe, he then warns us against allowing ourselves "to be led into anger and invective," and advises us to "pause a moment to reflect quietly, as a thinking creature should" (1, p 269) From Barbusse's quotations one can gather that, at least in a world safe with Socialism, anger and intelligence do not go together

In the relatively recent works of a neurologist, an individual psychologist, and a literary man, then, we find generalizations about

---

*Received in the Editorial Office on May 14, 1936

[1] One might ask, inferior in what, to whom, where?

anger which are in accord with the oft-quoted phrase, "Whom the gods destroy they first make mad." The foundation of those generalizations are not objective findings. One of the chief purposes of the present study is to consider with the use of objective data the extent to which, if any, these generalizations are warranted. The significance of this question for this study will be clearer if the presentation of it is preceded by a discussion of previous studies and a description of the methods used and general approaches made in the present one. We will next direct our attention to these aspects of the problem.

PREVIOUS STUDIES

Of the recent studies in anger treated on a psychosocial level (4, 7, 8, 10, 11), only the last two deal with aspects considered in this study, namely, the relationship with intelligence or achievement.

The earlier of these two studies, Morrison's (10), limited itself to a study of feebleminded individuals. The sources for the material she used were the records of behavior and observations made on mentally-deficient children. She classified her anger records into six gradations. These gradations ranged from one—no observable anger—to six—anger at the slightest provocation several times a day. Duration and intensity of anger were similarly scaled. She reports very high relationships, .79 to .90, between the various aspects of anger. In the feebleminded group of individuals, the correlation between intelligence and anger is relatively high, the average *r* being 65.

More closely related to our study is that made by G. M. Stratton (11), who used a group of college students. He had the subjects keep records of their anger for three days by checking on a prepared list of 19 situations arousing anger and six possible effects of its manifestation. By the use of this method he obtained a correlation of — 102±.106 between intellectual work as measured by grades and anger obtained as described. His subjects were 371 college students.

METHODS AND MEASURES USED IN THE PRESENT STUDY

The method used by Dr. Stratton readily enough lends itself to a statistical determination of an anger index. In the light of the findings reported by the present writer in an earlier study (8), the adequacy of the week-end sampling used by Stratton is very doubtful. The fact is, a relatively adequate sampling of college life must, as is shown in the earlier study, include days wherein activities participated in are dominantly social in nature as well as those dominantly intellectual, school days as well as weekends The method used in this, as well as in the earlier study by the present writer, is one that is definitely less constrained than are responses to a checked list and yields information in many ways as close to the life of the individual as that obtained by the diary or life history methods used in the investigation of personality traits by William Stern or Charlotte Buhler, or by some of the sociologists from the University of Chicago school

The method used for obtaining the facts of anger reported here has been called the controlled diary. This method was first used by G S. Gates (4). In this method the subjects are instructed to keep a diary of anger or extreme irritation for one week. They are also given specific instructions (*a*) to record the date, time, length of time before or after meals, duration, and intensity, (*b*) to rate on a fivefold scale each experience of anger, the condition of health, and degree of irritability; and (*c*) to describe the exact cause in each instance, the responses made, impulses experienced but not yielded to, and the after-effects In working up the materials, however, in this study even more than in the earlier one (8), the emphasis will be on the consideration of those aspects of anger adjustments which are more closely related to personality organization Because of this emphasis, little or no consideration will be given to variations in time of day and similar factors for which data are available[2] Instead, we will consider in some detail such factors as condition before anger situation arose, style of adaptation to reality, level of assertiveness, manner of expression, way of managing anger impulses, and organization of personality after anger reactions.

The subjects used in this study were 99 psychology students, mostly sophomores, with a sprinkling of juniors and seniors, at

---

[2] For facts concerning the duration of anger reactions in college students as well as daily variations, the reader is referred to the earlier study (8), as well as to (4) and (7)

Oregon State College The records of six of the students were not complete enough to be used, hence, with the exception of materials reported in Table 1, the data on anger are based on materials from 93 students The calculation of the data in Table 1 was not included in the original plans for working up the data, and, when in preparation of this article it was decided to report them, the record-book containing the intelligence and achievement scores was no longer available As a consequence, the correlation and frequency figures were made on only those anger charts which contained records of such scores On intelligence there were 78 such cases, on achievement scores, 90.

TABLE 1

RELATIONSHIPS WITH FREQUENCY OF ANGER RESPONSES

	r		Mean	S D $_{diff}$	S D $_m$
Intelligence	02± 08	S	4 77	2 68	53
		M	3 95	2 97	58
		I	5 04	3 09	61
Achievement	— 07± 07	S	4 67	3 08	56
		M	4 13	2 63	48
		I	4 40	2 93	53

Intelligence[3] in this study was based on scores made on the Otis *Self-Administering Examination, Higher, Form A* On a 20-minute time limit the scores ranged from 21 to 62 The total group was divided into three sub-groups. The range of test scores was 21 to 38 for the inferior group, 39 to 46 for the mediocre group, and 47 to 62 for the superior.

Achievement[3] as used in this study was based on six objective tests in psychology [4] The scores on the six tests, given at well-spaced intervals during the school session, ranged from 206 to 363 The range of scores for the superior, mediocre, and inferior divisions in achievement were, respectively, 206 to 260, 261 to 293, and 294 to 363

ANGER AS A PROBLEM IN PERSONALITY ADJUSTMENTS

A concrete picture of the nature of the data as well as the nature of anger reactions evoked can be obtained by a study of the anger charts of selected individuals For this purpose, the following charts of the brightest and dullest males in the group are presented.

---

[3] Intelligence in the tables is symbolized by *In* and Achievement by *Ac*

[4] Actual samplings of the tests used can be seen in (9)

ANGER CHART OF THE BRIGHTEST MALE STUDENT
(Otis Score—61, Achievement Score—308)

Cause	*Impulse controlled*	*Impulse yielded to*	*Effect*
Large assignment in Vocational Education course	To make sarcastic remark to teacher		Defiance toward the course
Editorial in Stanford paper		Commented about it to colleague	Laughed about the incident
Examination	Tear up questions	Sigh of anguish	Nervousness
Recalled forgotten appointment	Call up and apologize	Slight feeling of fear	Decision to try to remember better
Poor grade on examination	Razz instructor on type of exam	Made alibi	Rationalized
Awakened by noise	To throw vacuum sweeper downstairs	Cursed slightly, moved cover over head	Wished I could go some place and never be disturbed
Assignment of 140 pages discovered	To beat up an instructor	Swore	Trembling and excited feeling
More work than could complete	To throw book away		Feeling of hopelessness of studying anymore
Argument	To sit on him or trample	Raised voice and put up defense	Nervousness and inability to study and concentrate
Phone call	To cut wires	"Damn that phone"	
First whistle	To miss class and sleep	Reminded that I would miss breakfast	Realized I would be uncomfortable all morning

ANGER CHART OF THE DULLEST MALE STUDENT
(Otis Score—21, Achievement Score—260)

Cause	*Impulse controlled*	*Impulse yielded to*	*Effect*
Short in exam	Let matter go	Stuck to it	Discouraged
Unable to find assigned material	To read light fiction or play golf	Leave it go for awhile	Excited
Interrupted speech	To stop practicing	Asked others to mind own business	Tired, nervous
Couldn't finish study for speech	To let it go	Got angry	Relaxed, went to bed
Stumbled, fell in mud		Couldn't understand how it happened	Humor
Letter asking financial assistance			Indifference

A glance at the charts clearly indicates anger reactions to be related to planned fulfillments, experienced frustrations, and reactions thereto—typical problems of personality adjustments For example, consider the dullest individual His Otis score of 21 in 20 minutes is 11 points below that which indicates average intelligence and 21 points below that which indicates average intelligence and 21 points below that which signifies average college intelligence In spite of this relatively low level of mentality he succeeds in making an average grade in a course. Examine his anger chart: The sources of irritation for his anger reactions were most of the time of an intellectual nature Of the six experiences he reports, four are intellectual, one is accidental, and one is financial The accident, stepping into a mud puddle, evoked humor The effect of the financial stimulus was indifference. But the effects of the irritations brought on by stimuli which were intellectual in nature were consistently negative or dissociative, discouragement, excitement, tiredness, nervousness, sleep.

A glance at the anger reactions of the brightest individual gives a completely different picture of emphasis To begin with his experiences are more varied, both situationally and emotionally considered For example, liberally interpreted, 6 of his 11 anger experiences of the week may be said to be in response to intellectual challenges. But aside from the fact that the bright individual is carrying a heavier curriculum load, one of his six experiences classified as intellectual is in response to an extracurricular activity—a reaction to an editorial in a paper from a competing university Four of the 11 experiences are irritations in response to situations which are dominantly personal or social in nature A wider range of responsiveness is indicated by the fact that his reactions range from laughing to intense irritability His richer style of assertiveness is indicated by the inclusion in his repertoire of anger, reactions which are defiant, rationalized or realistic as well as dissociative in the sense of evoking discouragement, failure, or restlessness If one knew nothing more about these two individuals than their anger charts, one could easily enough guess which is the richer personality, which one is more involved in assertive reactions in campus life in terms of the nature of planned fulfillments and frustrations

In discussing the story of mankind Barnes (2, p v) advises us that

> In some epochs the mere struggle for existence has been most important. In others religion has loomed largest in the interest of peoples In another age the struggle for wealth and power has been dominant Among some peoples, art and things of beauty have seemed to possess most significance When treating a particular people in any era, I have attempted to stress especially the most characteristic and illuminating traits of that age and civilization

In dealing with the anger adjustments of individuals a similar approach seems advisable This is in accord with the general conclusion of our earlier study to the effect that strictly speaking there is no such thing as anger and that realistically considered anger is a problem in psychosocial dynamics This conception is contrary to the McDougall notion of a list of emotions including anger accompanied by a list of instincts It is also contrary to the conception of rage as an innate pattern [5]

---

[5] A notion associated with the name Watson and also explicity expressed by Burt (3, p 416) "Not only is the combative instinct itself inherited, but differences in its strength seem demonstrably inherited, too "

Some of the general findings of the earlier study which warrant *this conclusion are* The causes or stimuli which arouse anger in the individuals investigated range all the way from a thwarting of a desire to do nothing to the interference with a desire to do everything. The impulses felt during anger may be dominantly verbal, physiological, psychological, or social in nature and may range from a pleasant reply to doing violence to the offender The after-effects range from a very reduced self-feeling to feelings of exalted self-importance The organization of home-life and geographic setting play large rôles in determining the time of anger occurrences and the sex of the individual influences the sources of irritation

In the light of these findings it seems clear that an understanding of anger adjustments of any given individual calls for a consideration of the life history of that individual as well as the psycho-social factors which influence his anger behavior A configurational emphasis on what is relevant in a given situation and the extent to which it is relevant, not a predetermined list of innate or acquired patterns, seems to be the most promising as well as realistic approach to the study of anger adjustments in humans

FINDINGS AND INTERPRETATION

1. *Relationships of Intelligence and Achievement with Frequency of Anger Responses* The facts of relationship which exist between frequency of anger responses and intelligence and achievement are indicated in two ways (1) by correlations for comparison with Stratton's data (11), and (2) by comparable measures of central tendency and variability for the superior, mediocre, and inferior individuals in the group, considered in terms of intelligence as well as achievement

The correlations found here are quite in accord with those reported by Stratton (11). The latter investigator, without reporting data other than the correlation between intellectual work and anger, goes into a relatively lengthy discussion which, with reservations, favors the conclusion that the greater impulse to get angry is an impediment to scholarly achievement In the present study, no attempt will be made to build up an argument about the meanings or possible meanings of correlations which, as in this instance, are close to zero Instead, the group has been divided into three sub-groups for the purpose of searching out differences in frequency of anger and other related factors that are not revealed by the correlation method

As treated, the results indicate that individuals who are mediocre in either intelligence or achievement average less anger responses than do either the superior or inferior individuals The difference between individuals who are superior in intelligence and those who are mediocre is 81 in favor of the former The S. D of the difference is 79 and the Diff /S. D diff is 1 03 The difference between the inferior and mediocre is 1 08 in favor of the former The S D is 84 and the Diff /S D diff is 1 29 Expressed in terms of percentages, this means that the intellectually superior as compared with the mediocre average 20 per cent more anger responses, and the inferior 27 per cent more Considered in terms of achievement, the difference in means between the superior and the mediocre is 54, the S D of the difference 74, and the Diff /S D diff 73 The difference between the mediocre and the inferior is 27 in favor of the latter, the S. D. of the difference 72, and the Diff /S D. diff 38 This, expressed in terms of percentages means that, as compared with the mediocre in achievement, the superiors average 13 per cent more anger responses and the inferiors 7 per cent more.

Comparison between the differences indicates that, on frequency of anger responses in this group, intelligence differentiates the mediocre from the inferior and superior more than does achievement

2 *Relationships with Condition before Anger Situation Arose.* Burt (3) talks of anger as being tied up very closely with disturbed condition of the body "An irritable frame of mind," according to him, "rests often on a disturbed condition of the body, and the milder forms of anger—petulance, peevishness, and sensitive fretfulness—appear, in a majority of cases, to be the indirect outcome of physical weakness or physical irritation " In this connection, Burt also quotes Healy as saying that he has failed to find excessive irritability without some physical cause. Statements like the foregoing suggest a close connection between physical conditions of an individual and susceptibility to irritation In Table 2 are reported the physical conditions in which our subjects were before the anger situation arose The many descriptions given by the subjects concerning their conditions were divided into three categories. Disequilibrium, equilibrium, and asleep Un-

der "disequilibrium" are included such descriptions as the following: tired, sleepy, tired and hungry, hungry, headache, cold, irritable, worried, eyes tired, disappointed, discouraged, disgusted, nervous, in great hurry, restless, despondent, blue, exhausted, lazy, bored. Under "equilibrium" were included high spirits, gay, happy, good, relaxed, contented, congenial, normal, O. K. The caption "asleep" needs no explaining. When no conditions were described in this study, the self-rating of the individual on his condition was used to determine whether it represented a state of equilibrium or disequilibrium. The data are reported in terms of percentage of experiences classified as indicated in the table.

TABLE 2

Relationships with Condition before Anger Situation Arose*

	Disequilibrium			Equilibrium			Asleep		
	S	M	I	S	M	I	S	M	I
In	77.78	86.49	80.00	22.22	12.16	10.00	0	1.35	10.00
Ac	81.39	76.00	79.25	16.28	19.00	20.75	2.33	5.00	0

*The meanings of abbreviations used in the tables are: *In*, intelligence, *Ac*, achievement; *S*, superior, *M*, mediocre, *I*, inferior.

The data reported in the foregoing table indicate that the differences between superior, mediocre, and inferior individuals are larger in intelligence than they are in achievement. Considered in terms of intelligence, the mediocre are the most frequently found in a state of disequilibrium, and the superior least often so found. The superior group is more often in a state of equilibrium and the inferior least, and the inferior group in intelligence is by far more often in a state of sleep. Considered in terms of achievement, the mediocre are the least often in a state of disequilibrium before the anger situation arises, and the superior most often. And unlike the differences found in intelligence, in achievement it is the mediocre group that is found most often in a state of sleep.

In interpreting the foregoing data, it is perhaps well to keep in mind that the superior in intelligence, not those who are superior in achievement, are probably closer observers of their conditions before anger. In checking the descriptive nature of the responses, it was discovered that 62 per cent of the individuals who are rated superior in intelligence used descriptive phrases as well as a self-rating in evaluating their condition; only 49 per cent of the mediocre and 41 per cent of the inferior did On achievement from high to low, the percentages are 44, 51, and 53 Of all the conditions, "tired" is given most frequently by all individuals It is more often given by the mediocre in intelligence and achievement than by the superiors and inferiors. In the mediocre group tiredness is given as a descriptive phrase 31 times, in the bright group 21 times, and in the dull group 16 times In terms of achievement, these differences are, going from superior to inferior, 17, 39, and 21 "Sleepy" is the next most frequently reported condition, and on that condition again the mediocre report more frequently The differences in terms of frequency on intelligence are, going from high to low, 9, 13, and 2, respectively On achievement they are 8, 9, and 6 Grouped in terms of intelligence but not achievement, the superior more often described themselves as being in a state dscribed as "good," "gay," "normal," than do the other two groups This difference is indicated in the foregoing table by the fact that the superior report themselves to be in a state of equilibrium 22 22 per cent of the time as against 12 16 and 10 for the other two sub-groups

3 *Relationships with Nature of Anger Adjustments* In his *Fundamentals of Human Motivation,* Dr Troland (12, pp 385, 386) described what he considers to be the correlates of inferiority-superiority motives in relation to anger By and large, the general impression one gets from reading his involved language is that it is his opinion that, when a man of low intelligence is insulted, he more likely will respond with a physical attack upon the offender On the other hand, in a man of higher intellectual development, according to Dr Troland, "the instrumental images may be visual or auditory, as well as kinaesthetic," and because of this he will less likely attack the offender and more likely will, perceiving an attack to be an inferior reaction, retaliate in words or merely walk away

Burt comes to the same general conclusion as Dr Troland on the basis of a more evolutionary treatment of the topic of anger

> As all creatures but the lowest, according to him, are endowed with special organs of sex, so many of the higher are equipped with special organs for defense and for aggression—the bull with horns, the fighting cock with spurs, the tiger with tooth and claw Man, when fully grown, being—in the defi-

> nition of Herr Teufelsdrockh—"a tool-using animal with the gift of speech," fights now with artificial instruments of violence and now with his tongue (3, p 416)

On actual objective findings on an anger study in the feeble-minded group, Morrison (10) reports a difference in the manner of expression in the idiots, imbeciles, and morons. On the lowest level, the causes of anger are chiefly bodily conditions, interference with movement. In the imbecile group, the causes are chiefly the thwarting of purposes or interference with work. In the higher level of mental deficiency, she reports causes of a personal-social nature, such as quarrels.

Are there any differences in the manner or form of expression between superior, mediocre, and inferior college students? The findings concerning the nature of types of adjustments made by the various individuals in anger situations will be presented in three ways. (*a*) styles of adaptation to reality, as classified in Table 3,

TABLE 3

RELATIONSHIPS WITH STYLE OF ADAPTATION TO REALITY

	Faces total situation			Faces partial situation			Compromise			Non-adaptive		
	S	M	I	S	M	I	S	M	I	S	M	I
In	1.05	0	0	28.42	28.03	33.33	15.79	19.75	25.49	54.74	52.22	41.18
Ac	.95	0	0.92	21.90	32.83	40.37	22.86	21.21	12.84	54.29	45.96	45.87

TABLE 4

RELATIONSHIPS WITH LEVEL OF ASSERTIVENESS

	Aggressive			Defensive			Submissive		
	S	M	I	S	M	I	S	M	I
In	29.47	26.58	31.57	52.63	48.74	49.02	17.90	24.68	19.61
Ac	36.19	27.78	32.11	47.62	50.00	43.12	16.19	22.22	24.77

TABLE 5

RELATIONSHIPS WITH MANNER OF EXPRESSION

	Verbal			Action			Physiological			Inert		
	S	M	I	S	M	I	S	M	I	S	M	I
In	52.63	49.37	59.80	24.21	20.88	21.57	6.32	10.76	4.90	16.84	18.99	13.73
Ac	51.43	55.05	43.12	24.76	21.21	28.44	9.32	4.55	13.76	14.29	19.19	14.68

(*b*) levels of assertiveness—aggressive, defensive, or submissive as indicated in Table 4, and (*c*) manner of expression, sub-divided in terms of verbal, action, physiological, and inert reactions as presented in Table 5 The results in Tables 3, 4 and 5 are all reported in percentages

A survey of the data presented in Table 3 suggests that by and large the anger reactions of individuals from all levels of intelligence and achievement are largely non-adaptive Only one reaction (1 05 per cent) is classified as representing a facing of total reality, and that was manifested by an individual of superior intelligence No such reactions are found in either the mediocre or inferior groups The inferior in intelligence average a larger percentage of reactions that are classified as partial facing of reality and also average a larger percentage of reactions which represent compromises with realty The superior average more non-adaptive responses Other more specific differences between the three groups on the two factors can be obtained from reading Table 3

The intellectually mediocre average less aggressive and defensive reactions than do the inferior or superior, and more submissive reactions than the other two. The intellectually superior average more than the other two groups only in the percentage of defensive reactions Considered in terms of achievement, the superiors average most aggressive reactions, are middling in defensive reactions, and average the smallest percentage of submissive reactions. The mediocre in achievement average most defensive reactions, middling in submissive reactions, and lowest in aggressive reactions

Whatever the differences may be when considered in the evolutionary manner of Burt or in the objective findings of various levels of the feebleminded group concerning the manner of expressing anger, in a group composed of college students, regardless of level of intelligence, no such consistent differences are found The specific differences are reported in Table 5 As indicated in this table, the intellectually superior average the highest percentage of action responses, average middling on physiological reactions and inert reactions, and are below the inferior, though above the mediocre, in the percentage of verbal reactions. The superior in achievement are lowest in inert reactions and middling in verbal reactions, action responses, and physiological reactions

Of the many specific impulses felt during the experience of anger, to make a verbal retort was, with but one exception, reported most frequently by persons from all levels of intelligence and achievement. The exception was the impulse to injure inanimate objects reported by the brightest group. Expressed in terms of percentages, the intellectually superior average 24.18 per cent verbal retorts, the mediocre 24.36, and the inferior 19.59. Considered in terms of achievement, superiors average 20 per cent such responses, the mediocre 18.72 per cent, and the inferior 20.41 per cent such responses. Contrary to the implications of the quotations enclosed in this section, the next most frequently reported single felt impulse, with the exception mentioned, was to do injury to the offender. In that, the superior intellectually average more responses than do either of the other two. The percentages of reactions of this type expressed by the intellectually superior were 18.68, by the mediocre 8.97, and by the inferior 12.37. Considered in terms of achievement, the story is somewhat different. Only 12 per cent of the superiors expressed an impulse to do physical injury to the offender, 11.76 of the mediocre, and 19.39 of the inferior. It should be remembered, however, that this is merely a description of the impulses felt, rather than those released.

To run away and leave the room, which, according to Troland's description, would represent a high level reaction, was manifested most often by the mediocre group. Four such reactions, 4.3 per cent were listed by the intellectually superior, 9 or 5.77 per cent by the mediocre, and 2 or 2.06 per cent by the inferior. Considered in terms of achievement, again the story is different. Here the superiors do average the highest number of such reactions, 5 per cent, the mediocre average 2.67 per cent and the inferior 4.08 per cent. The specific-felt impulse of retaliation in kind was least often expressed by the intellectually superior. The percentages from high to low are, respectively, 1.10, 2.56, and 8.25. Considered in terms of achievement, the percentages are 4.0 for the superior, 3.21 for the mediocre, and none for the inferior. The specific response "to blame someone else" was found once among the mediocre in intelligence and once among the inferior, and was absent in the intellectually superior group. In terms of achievement, the percentages are 2 for the superior, 54 for the mediocre, and 1.02 for the inferior. "Razzing the instructor" was expressed by only one individual, who was superior both in intelligence and achievement. A desire to invent "Bigger and better cuss words" was expressed by only one individual, superior in intelligence and achievement. To "dash for home" was expressed only once, by one mediocre and one inferior individual, and not at all by the superior group.

4. *Relationships with Management of Impulse.* What is the relationship between the manner in which an anger impulse is managed with intelligence or achievement? To get a more specific picture of this than is obtained by the foregoing more general descriptions, the responses made and the impulses felt in each instance were classified in terms of a fourfold division of impulse-response configurations. The divisions, as indicated in Table 6, are impulse inhibited, impulse partially inhibited, impulse yielded to, and impulse better than response. The results are reported in percentages.

TABLE 6
RELATIONSHIPS WITH MANAGEMENT OF IMPULSES

	Impulse inhibited			Impulse partially inhibited			Impulse yielded to			Impulse better than response		
	S	M	I	S	M	I	S	M	I	S	M	I
In	26.51	32.14	44.29	61.44	27.96	56.99	12.05	23.57	10.75	0	0	4.30
Ac	31.25	35.52	25.53	54.17	45.36	60.64	14.58	16.39	13.83	0	2.73	0

As classified, the largest per cent of impulse-response configurations fall into the partially inhibited division, regardless of level of intelligence or achievement. Impulse inhibited averages the next most categories from all levels, and the impulse better than response category the least. The intellectually superior rank highest in the impulse partially inhibited. The mediocre average most in impulses yielded to, and the inferior average most in impulses inhibited and impulses better than response. Considered in terms of achievement, the mediocre group averages highest in impulse inhibited, impulse yielded to, and impulse better than response. The inferior in achievement average highest in impulse partially inhibited.

Another bit of insight can be gained concerning the nature of the significance of the way in which the impulse was managed by evaluating the after-effects in terms of level of socialization. For the purpose, the effects were classified on a fourfold scale labelled: individual, individual-social, social-individual, and social. So classified, the percentages for the intellectually superior on the four categories are: 45 individual, 18 individual-social, 9 social-individual, and 28 social. Comparable percentages for the mediocre are 37, 13, 10, and 40, and for the inferior, 32, 20, 8, and 40. Considered in terms of achievement, the percentages for the superior are: 39 individual, 18 individual-social, 12 social-individual, and 32 social. Respective percentages for the mediocre are 41, 11, 7, and 41, and for the inferior, 39, 22, 9, and 30.

A glance at these percentages suggests that there is a consistent and well-directed tendency for the intellectually superior to *manifest the largest percentage of individual reactions and the* smallest percentage of social reactions. No regular tendency appears in the groups considered in terms of achievement. The mediocre average most social responses as well as most individual reactions.

5. *Relationships with Organization of Personality after Anger Reactions.* Do anger responses tend to organize or disorganize the personality? Are there differences in the extent to which they tend to organize or disorganize in individuals of various levels of intelligence or achievement? In Table 7 are presented the reactions after being classified into the general categories of the reactions which express dissociation, integration, or no reaction. On the dissociative trend we include such specific after-effects as irritable, sorry for self, *disgusted, discouraged, ashamed, trembling, forgetting to eat,* distracted, sad, hurt, nervous, increased dislike for person causing anger. An after-effect was classified as symptomatic of an integrative trend if balance or expansion of the ego is indicated. Illustrative of such after-effects are satisfaction, triumph, feeling of energy, calm, O. K., relief, efforts continued. The third caption used is self-explanatory. The results are reported as percentages.

TABLE 7

RELATIONSHIPS WITH ORGANIZATION OF PERSONALITY AFTER ANGER REACTIONS

	Dissociative trend			Integrative trend			No reaction		
	S	M	I	S	M	I	S	M	I
In	64.45	61.44	65.63	13.33	9.15	11.46	22.22	29.41	22.91
Ac	63.27	63.98	61.00	6.12	11.29	19.00	30.61	24.73	20.00

The data presented in Table 7 indicate that, for individuals of all levels of intelligence and achievement, anger responses are far more frequently dissociative in effect The relative differences within the three groups considered in terms of intelligence or achievement under dissociation are small In the percentage of effects which lead towards integration, the superior group intellectually rate highest and the mediocre lowest Considered in terms of achievement, it is the inferior who average the largest percentage of reactions that indicate an integrative trend

GENERAL SUMMARY

1 *Approach* Anger is here presented as a problem in personality adjustments which can be understood only in the light of a knowledge of the life history of the individual who is manifesting the anger and all the psychosocial factors which are influencing him at the time of the anger response The meaning of anger as a problem in personality adjustment is illustrated with the anger charts of the brightest and the dullest males in the group

2 *Method* The method used in this study was the controlled diary, wherein were obtained facts from which could be gathered information concerning such aspects of personality organization that may be related to anger, as condition before anger situations arise, style of adaptation to reality, level of assertiveness, manner of expression, management of impulses, and organization of personality after anger reactions Data are presented in terms of differences for superior, mediocre, and inferior individuals within the group in intelligence as measured by the Otis *Higher Examination* and achievement as measured by six objective tests

SUMMARY OF FINDINGS

1 *Relationships of Intelligence and Achievement with Frequency of Anger Responses* In a feebleminded group anger and intelligence are closely related The correlation of these factors in such a group is 65 Imbeciles definitely are more readily susceptible to anger than idiots, and morons more so than imbeciles But in a college group where the individuals studied range in intelligence from low average to near-genius, the correlation is practically zero In the present study, the frequency of anger correlates .02± 08 with

intelligence and —.07±.07 with achievement Intellectually superior individuals in the college group average 20 per cent more anger responses than the mediocre, and the inferiors average 27 per cent more. As compared with the mediocre in achievement, the superiors average 13 per cent more and the inferiors 7 per cent more These differences are presented in terms of means and S. D.'s

2. *Relationships with Condition before Anger Situation Arose* Differences between superior, mediocre, and inferior individuals are larger in intelligence than achievement, as far as condition before the anger situation arises is concerned Considered in terms of intelligence, the mediocre individuals are most often found in a state of disequilibrium, and the superiors least often so found The superior group is more often in a state of equilibrium and the inferior least, and the inferior group in intelligence is by far more often in a state of sleep. Considered in terms of achievement, the mediocre are the least often in a state of disequilibrium before the anger situation arises, and the superior most often. The superior intellectually most often describe themselves as being in a state of equilibrium

3 *Relationships with Nature of Anger Adjustments*

a. Style of adaptation to reality. Anger reactions of individuals from all levels of intelligence and achievement are largely non-adaptive. Specific differences on a number of reactions which represent a facing of reality, partial adaptation to reality, compromise with it, as well as non-adaptiveness, may be found in the body of this article.

b Level of assertiveness. The intellectually mediocre average less aggressive and defensive reactions than do the inferior or superior, and more submissive reactions than the other two The intellectually superior average more than the other two groups only in the percentage of defensive reactions Considered in terms of achievement, the superiors average most aggressive reactions, are middling in defensive reactions, and average the least percentage of submissive reactions. The mediocre in achievement average most defensive reactions, middling in submissive reactions, and lowest in aggressive reactions.

c Manner of expression Unlike generalizations found in literature or findings in the feebleminded group reported by Morrison, there are no large differences in manner of expressing anger in a group of college students of various levels of intellect or attainment Some of the more interesting differences reported in Table 5 are The intellectually superior average the highest percentage of action responses, middling in physiological reactions and inert reactions, and are second to the inferior group in the percentage of verbal reactions. The superior in achievement are highest in none of the categories, lowest in inert reactions, and middling in verbal reactions, action responses, and physiological reactions The differences in some specific reactions, like verbal retort, doing physical injury to the offender, running away and leaving the room, are reported in the body of the paper

4 *Relationships with Management of Impulses* When the impulses felt and the responses made are classified in impulse-response configurations, the largest number of impulse-response patterns fall into the partially inhibited division, regardless of level of intelligence or achievement Impulse inhibited averages the next most categories from all levels, and the impulse better than response category the least. The intellectually superior rank highest in the impulses partially inhibited The mediocre average most in impulses yielded to, and the inferior average most in impulses inhibited and impulses better than response Considered in terms of achievement, the mediocre group averages highest in impulse inhibited, impulse yielded to, and impulse better than response The inferior in achievement average highest in impulse partially inhibited

Classified on a fourfold category ranging from individual to social, the reactions of the intellectually superior are found to be most individual, and those of the intellectually inferior most social.

5 *Relationships with Organization of Personality after Anger Reactions* The after-effect responses were classified as indicating a trend towards dissociation or integration The relative differences between the three groups considered in terms of intelligence or achievement under dissociation are small; anger responses are far most frequently dissociative in effect for individuals of all levels of intelligence or achievement In the percentage of effects which lead towards integration, the superior intellectually rate highest and the mediocre lowest Considered in terms of achievement, it is the inferior who average the largest number of reactions that indicate an integrative trend

REFERENCES

1 Barbusse, H Stalin New York Macmillan, 1935

2 Barnes, H E The history of western civilization (2 vols) New York Harcourt, Brace, 1935.

3 Burt, C The young delinquent New York Appleton, 1925 Pp 619

4 Gates, G S An observational study of anger *J Exper Psychol,* 1926, 9, 325-331

5 Herrick, C J The thinking machine Chicago Univ Chicago Press, 1929 Pp xii + 374

6 Lazell, E W The anatomy of emotion man's two natures New York Century, 1929

7 Longworthy, A H A comparative study of anger and irritation in children and adults. Mount Holyoke College Dissertation Study, June, 1935

8 Meltzer, H Students' adjustments in anger *J Soc Psychol,* 1933, 4, 281-309

9 Meltzer, H , & Bailor, E. M Developed lessons in psychology New York Harcourt, Brace, 1929 Pp x + 199

10 Morrison, B M. *A study of the major emotions in persons of defective intelligence* *Univ of Calif Publ Psychol,* 1924, 3, No 3, 73-145

11. Stratton, G M *Anger and fear their probable relation to each* other, to intellectual work, and to primogeniture *Amer J Psychol,* 39, 125-141

12 Troland, L T The fundamentals of human motivation New York Van Nostrand, 1928 Pp. xiv + 521

Psychological Service Center
St Louis, Missouri

CHANGES IN SEX DIFFERENCES IN PROBLEMS AND INTERESTS OF ADOLESCENTS WITH INCREASING AGE*

Teachers College, Columbia University

PERCIVAL M SYMONDS

Do boys and girls become more or less alike in their problems and interests as they grow older? Discussions of sex differences in psychological literature are largely concerned with physical and intellectual differences, and accurate information regarding differences in interests is scanty One may easily observe that boys and girls, before the age, say, of 10, play together, in the home and in nursery school, kindergarten, and lower grades of the elementary school While girls may prefer dolls and boys may prefer trains as toys, they readily join in the same games Before adolescence, however, this association disappears and when it reappears in adolescence it is on a wholly different basis as determined by the maturing sex functions Boys and girls at all periods of adolescence have different interests in reading (5, 7), writing (2), movies (11), radio programs (4), games (3, 8, 9, 14), conversation (15, 16), and vocational selection (1, 6, 10, 12) Do these differences become more or less intense?

A survey of the problems and interests of adolescents and adults furnishes data which throw light on this problem Fifteen areas of human interest (briefly noted in the tables accompanying this paper) were ranked by several hundred high school pupils, college students, and graduate students in education in order of being personal problems, and in order of being of interest for reading or discussion. By taking average ranks for the various groups it is possible to note differences by sex for different maturity levels Tables 2, 3, and 4 show these differences for 276 college men and 387 college women, also differences for 784 high school boys and 857 high school girls, and 73 men and 111 women who are graduate students of education

The outstanding generalizations from these data are that there is more difference in general between adolescent boys and girls in

---

*Received in the Editorial Office on May 19, 1936

TABLE 1

SUMS OF DIFFERENCES IN AVERAGE RANKING BETWEEN MEN AND WOMEN

	Problems	Interests
High school group	47	78
College group	72	112
Adult graduate group	56	51

TABLE 2

AVERAGE RANKINGS FOR 15 ITEMS FOR PROBLEMS AND INTERESTS OF 147 STUDENTS AT PURDUE UNIVERSITY AND 516 STUDENTS AT KANSAS STATE TEACHERS COLLEGE, EMPORIA, KANSAS

		Boys		Girls	
		Problems	Interests	Problems	Interests
1.	Health	6.81	5.93	7.58	8.03
2.	Sex	8.67	5.96	9.76	7.21
3.	Safety	10.41	10.44	11.38	12.23
4.	Money	4.07	7.01	5.71	9.15
5.	Mental health	8.16	7.88	7.02	8.03
6.	Study	6.08	9.17	5.65	9.06
7.	Recreation	9.12	7.06	9.56	7.84
8.	Personal qualities	7.42	6.95	7.43	6.08
9.	Family relationships	9.83	9.87	9.43	8.85
10.	Manners	8.96	7.66	8.78	6.46
11.	Personal attractiveness	7.96	8.31	7.01	5.87
12.	Daily schedule	8.29	11.14	7.52	10.14
13.	Civic	8.80	8.75	7.96	8.59
14.	Getting along with others	8.82	7.89	9.14	7.24
15.	Philosophy	7.28	5.95	6.35	4.80

their *interests* than in their *problems* and there are more differences between college men and women than between high school boys and girls, and adult men and women.

Evidence for these two generalizations may be seen in the sums of differences (taken without regard to sign) in ranking between men and women for the three groups (Table 1).

In both the high school and college groups it will be seen that the difference between the sexes in interest is greater than the difference in problems.

The data also indicate that men and women are most unlike during the period of later adolescence. At the time when all the powers, physical and mental, have matured to their maximum and when the process of mating is beginning to be the main life concern, men

TABLE 3

DIFFERENCES BETWEEN AVERAGE RANKINGS OF 276 MEN AND 387 WOMEN OF 15 ITEMS FOR PROBLEMS AND INTERESTS FOR 147 STUDENTS AT PURDUE UNIVERSITY AND 516 STUDENTS AT KANSAS STATE TEACHERS COLLEGE, EMPORIA, KANSAS

A plus sign (+) indicates a higher ranking for men

		Difference for problems	$\frac{\text{Diff}}{\sigma_{diff}}$	Difference for interests	$\frac{\text{Diff}}{\sigma_{diff}}$
1	Health	+ .77	+2.2	+2.10	+6.7
2	Sex	+1.09	+3.3	+1.25	+3.6
3	Safety	+ .97	+3.5	+1.79	+7.0
4	Money	+1.64	+5.5	+2.14	+7.0
5	Mental health	−1.14	−3.6	+ .15	+ .5
6.	Study	− .44	−1.4	− .12	− .3
7	Recreation	+ .44	+1.5	+ .78	+2.3
8	Personal qualities	+ .01	0	− .87	−3.3
9.	Family relationships	− .40	−1.3	−1.02	−3.7
10	Manners	− .18	−0.7	−1.20	−4.3
11	Personal attractiveness	− .95	−3.5	−2.44	−8.7
12	Daily schedule	− .76	−2.2	−1.00	−3.7
13.	Civic	− .83	−2.8	− .16	− .5
14.	Getting along with others	+ .32	+1.0	− .65	−2.1
15	Philosophy	− .93	−2.9	−1.15	−3.3

and women are most different from each other. This differentiation proceeds from birth, but receives an acceleration at the time of adolescence. In later years as sex becomes less of a concern the differences in interests between men and women become less and probably vanish as senescence is reached.

An inspection of some of the specific items will indicate in what ways men and women are most unlike at the college period. The maximum difference is in interest in *personal attractiveness*, women being more interested in this item to the extent of 2.44 rank places. Even so, the difference is about the same or even less than at the high school age when the obtained difference is 2.72.

Opposed to this is the interest in *money* which is 2.14 ranks greater for men than for women. This states for us the familiar fact that at this age men are interested in establishing themselves vocationally and setting up their competence, while women apparently are more interested in the process of sexual selection.

TABLE 4

DIFFERENCES BETWEEN AVERAGE RANKINGS OF 784 BOYS AND 857 GIRLS IN HIGH SCHOOL AND ALSO 73 MEN AND 111 WOMEN GRADUATE STUDENTS OF EDUCATION FOR 15 ITEMS REPORTING AREAS OF HUMAN PROBLEMS AND INTERESTS

A plus sign (+) indicates a higher average ranking for boys than for girls

		High school		Adults	
		Difference for problems	Difference for interests	Difference for problems	Difference for interests
1	Health	— 11	+ 96	+ 64	+ 29
2	Sex	+ 24	+ .16	+1 20	+1 67
3	Safety	+ 54	+1.39	— 31	+ 34
4	Money	+ 58	+ 93	+ 86	+ 46
5.	Mental health	— 47	+ .26	— 16	— 19
6	Study	+ 54	+ 62	— 04	— .13
7	Recreation	+ 07	+ 74	+ 20	+ 10
8	Personal qualities	+ 14	— .17	+ 35	+ .44
9	Family relationships	— 02	— 18	+ 30	+ .28
10	Manners	—1 11	—1 16	+ 51	+ 25
11.	Personal attractiveness	—1 70	—2 72	—1 39	—1 76
12.	Daily schedule	+ 44	— 09	—2 00	— 68
13	Civic	+ 47	+ 77	— 26	— 27
14	Getting along with others	— 50	—1 13	— 05	— 76
15	Philosophy	— 11	— 36	+ 10	— 02

The greater interest of men in *health* is less obvious in everyday life, but this same difference in varying degrees has been noted in every school in which these data have been gathered and at all ages Apparently for men this is a variant of the drive for success, while for women the narcissistic tendency is more toward attractiveness than toward efficiency.

Men rank their interest in *sex* higher than do women. This should also be compared with the interest in personal attractiveness, and shows the well-known difference between the sexes in their attitude and response toward sex. For men the drive is positive and direct Men have no hesitancy in expressing their interest in women or in a woman and in pursuing the object of their interest and affection with ardor Women, on the other hand, while no less backward in making themselves attractive and in encouraging male advances, do not openly express as direct an interest in sex, but invite attentions from men.

That men are more interested in *safety* than women is not to be wondered at perhaps, since men are in general more active and have to participate in more dangerous undertakings and occupations I am inclined to believe, on the other hand, that women are not so sensitive to danger as men and that they accept dangerous situations with more hardihood and endurance The fact that women must go through the rigors of childbearing perhaps tempers their whole attitude toward the dangerous in life.

Men are less interested, at this age, in the *philosophy of life* than women, which may check with the common belief that women are more religious than men. My data show that this difference in less degree is present in high school boys and girls, but practically fades out with maturity. Quite possibly the greater activity and outwardness of men, particularly at the time when they are zealous for success in school and college, give them less time and interest for mental speculation, while the natural greater passivity of women makes them more receptive to speculations of a more general character

Two other differences may be commented on briefly One is the greater interest of women than men in *planning the daily schedule* High school youngsters find the daily schedule highly uninteresting—they are bored with sitting down and planning their day's activities Interest in planning comes earlier with women than with men, and the differences become still greater with adult women and men. This may be compared with the greater interest in the philosophy of life, and indicates that women are more ready than men to plan, to take stock of things, and to see the relationship of what they do with the larger issues of life. This points to an essential conservative character of women.

The other difference to be commented on is the fact that *mental health* is more of a problem for women than for men This harmonizes with the other differences already noted in that women are less active, more reflective, more religious, and accordingly tend to be more introspective It is well known that women are more prone to certain types of neuroses than men, and, as indicated by our data, they consider their mental health more of a problem

That women are more interested in *personal qualities, manners,* and *family relationships* is obvious, easily understandable, and fits in with other differences already commented upon, which need not be discussed in more detail here

SUMMARY

A comparison of the problems and interests of high school boys and girls indicates that the sex differences in *interests* are greater than the differences in *problems* at these levels.

These differences, both for problems and interests are greater in the period of late adolescence than at any other time of life At the time when men and women are most fecund biologically they are most unlike each other in their interests

When the differences are looked at in detail they are seen to agree with conventionally accepted descriptions of male and female characters. Men are more interested in *physical health, safety, money,* all representative of the urge toward success and personal establishment in society. Men are also more openly and directly interested in *sex* Women, on the other hand, are more passive and receptive. They differ most from men in their interest in *personal attractiveness* which represents their part in sexual selection

In addition, the greater passivity and receptivity of women makes them more introspective and more reflective This shows itself in their greater interest in a *personal philosophy,* in *planning the daily schedule,* and with their greater concern with problems of *mental health.*

Women also show their traditionally greater interest in people by their expressed greater interest in *manners, personal qualities,* and *home and family relationships*

REFERENCES

1 BEESON, M F, & TOPE, R E A study of vocational preferences of high school students *Voc Guid. Mag.,* 1928, 7, 115-119

2 COLEMAN, J H The written composition interests of junior and senior high school pupils. *Teach Coll. Contrib Educ,* 1931

3 CROSWELL, T R Amusements of Worcester school children. *Ped Sem,* 1899, 6, 314-371

4. EISENBURG, A Children and radio programs. Ph D dissertation, Teachers College, Columbia Univ, 1935.

5. HILDRETH, G H. Adolescent interests and abilities *J. Genet Psychol,* 1933, 43, 65-93.

6. HURLOCK, E B, & JANSING, C The vocational attitudes of boys and girls of high school age *J Genet Psychol*, 1934, **44**, 175-191

7 JORDON, A M Children's interests in reading Chapel Hill, No. Carolina Univ No Carolina Press, 1926

8. LEHMAN, H C, & WITTY, P A The psychology of play activities New York: Barnes, 1927 Pp xviii+242

9 MCGEE, Z A study of the play life of some South Carolina children *Ped Sem*, 1900, **7**, 459-478

10 MINER, J B An aid to the analysis of vocational interests *J Educ Psychol*, 1922, **5**, 311-323

11 MITCHELL, A M. Children and movies Chicago Univ Chicago Press, 1929

12 POULL, L E Interests in relation to intelligence *Ungraded*, 1922, 145-158, 176-192

13 TERMAN, L M., & LIMA, M Children's reading a guide for parents and teachers New York Appleton, 1926. Pp 363

14 TERMAN, L M, *et al* Genetic studies of genius Vol I Stanford Univ Press, 1925. Pp. xv+648

15. STOKE, S M, & WEST, E. D Sex differences in conversational interests. *J Soc. Psychol*, 1931, **2**, 120-126.

16 ——— The conversational interests of college students *School & Soc.*, 1930, **32**, 567-570

Teachers College
Columbia University
New York City

MONTH OF BIRTH AND AVERAGE INTELLIGENCE AMONG DIFFERENT ETHNIC GROUPS*

Teachers College, Columbia University

R Pintner and J B Maller

The problem of the relationship between the month of birth and the physical condition of the newborn and the related problem of seasonal variation in the growth of the child have been studied by many investigators Sanders (15) has summarized and discussed this field in detail He refers to at least 16 investigations on the average weight of infants born in different seasons of the year. Of these 16 reports, 13 indicate that children born in summer or in autumn are heavier than those born in the other seasons, two find no seasonal differences and one favors the winter months

These reports vary considerably with regard to number of cases and statistical procedure For example, Peller and Bass in 1924 reported on 2360 births The children born in summer showed an average weight of 80 to 87 gm greater than those born in winter Furthermore, 10 3 per cent of those born during July-September are below 2500 gm, whereas 15 3 per cent born during January-March vary below this weight These differences are said to be statistically reliable Similar findings are reported by Abels (1), Hansen (8), and Andersen (2), who report that children born in the fall were heavier than those born during the summer months

On the other hand, Brenton (6) in 1922 reporting on 2000 births finds no seasonal differences in weight Sanders in describing and commenting upon this study believes the data to show increase in weight in summer and autumn Sander's own conclusion, after discussing each of the 16 reports, is that on the whole there appears to be "some increase in birth size in late summer and fall, but it can hardly be considered significant" Reasons for this difference in weight suggested by several workers are, poor diet of the mother in winter, vitamin shortage in winter, increased sunshine in summer There is, however, no general agreement as to such causative fac-

---

*Received in the Editorial Office on June 5, 1936

tors Bakwin and Bakwin (4) and Hellmuth (9) found no constant seasonal variation in birth weight

In addition to such studies of weight at birth, Sanders reviews many others dealing with seasonal variations in *gains* in weight and height during infancy, the preschool period, and in later life The evidence, as a whole, is suggestive, but by no means conclusive that greater gains are made in summer rather than in winter Bakwin and Bakwin (3) report that, in certain anthropometric measurements, the differences were in favor of those born in autumn and winter. Causative factors suggested by various writers are temperature, nutrition, irradiation, endocrine activity, vitamin content of food at different seasons, seasonal morbidity, and the like When we turn from physical to psychological factors, we find very few studies dealing with the IQ of the child as related to the month of birth. Blonsky (5) in 1929 found the highest mean IQ for children born in spring Pintner (13) then reported on 4925 cases and found the lowest mean IQ for the winter months, January to March The greatest seasonal differences was only 1 7 IQ points, but it was statistically reliable The outstanding feature of these data was the large number of cases involved and the possibility of dividing the total cases into three different social groups as described in the study One factor, however, was not investigated, namely, the possible variations that might be caused by combining different ethnic groups It was suggested that ethnic groups might differ with reference to seasonal birth rates The combination of different ethnic groups differing in mean IQ might cause seasonal differences in the average IQ due to the variation in the proportion of births contributed by the various groups during any given season These considerations have led to the present study.

DESCRIPTION OF POPULATION

Our sample of cases consists of children in New York City public schools belonging to three ethnic groups, namely, Italian, Jewish, and Negro.[1] All the children tested were enrolled in the fifth grade of the elementary school, thus constituting an educationally homogeneous group The Negro children were pupils of schools in Central Harlem with an enrollment of 99 per cent Negro pupils. The Italian and Jewish children were pupils of schools in which the population was predominantly (85-100 per cent) Italian or Jewish, respectively

---

[1] For a description of differences between ethnic groups see (12)

MONTHLY BIRTH RATES OF ETHNIC GROUPS

We have studied the monthly birth rates for the different ethnic groups both for the group of children actually tested and for the city as a whole. Table 1 shows the distribution of our 6353 cases

TABLE 1

DISTRIBUTION OF FIFTH-GRADE PUPILS BY MONTH AND SEASON OF BIRTH FOR CHILDREN TESTED

	Italian		Jewish		Negro		Total	
Month of birth	Number	Ratio	Number	Ratio	Number	Ratio	Number	Ratio
January	188	1.10	254	.98	103	.90	545	1.01
February	164	1.06	230	.99	114	1.09	508	1.03
March	171	1.01	286	1.12	123	1.08	580	1.07
April	179	1.08	264	1.07	107	.97	550	1.05
May	172	1.01	269	1.05	123	1.08	564	1.04
June	158	.96	208	.84	86	.78	452	.87
July	164	.96	263	1.03	128	1.12	555	1.03
August	138	.81	283	1.11	116	1.01	537	.99
September	174	1.05	269	1.09	115	1.04	558	1.07
October	166	.97	261	1.02	114	1.00	541	1.00
November	172	1.04	204	.83	98	.86	474	.91
December	162	.95	208	.82	119	1.04	489	.91
Total	2008		2999		1346		6353	

	Italian		Jewish		Negro		Total	
Season of birth	Number	Ratio	Number	Ratio	Number	Ratio	Number	Ratio
Spring	522	1.03	819	1.08	353	1.04	1694	1.06
Summer	460	.91	754	1.00	330	.97	1544	.96
Autumn	512	1.02	734	.98	327	.97	1573	.99
Winter	514	1.04	692	.93	336	1.10	1542	.98
Total	2008		2999		1346		6353	

The table shows, for each ethnic group and for the total of all groups, the number of births in each month as well as the ratio of the obtained to the expected number In computing the expected number, the number of days in each month was taken into consideration An inspection of the table shows marked differences in number and ratio from month to month Thus, we note that in the Negro group the lowest number of births occurred in June

As will be shown later, the average IQ for June was higher than for any other month. Relatively low mean IQ's for the total group occur in December to May, inclusive. Relatively high mean IQ's occur from June to November. If now we calculate the proportion of births for the different groups for these two semi-annual intervals we have the following:

Period	Italian	Jewish	Negro	Total
June-November	48.5	49.6	48.8	49.6
December-May	51.5	50.4	51.2	50.4

Here it will be noted that the Italians and Negroes contribute a larger percentage of births to the December-May half year than does the Jewish group. It appears, then, that in our samples of the three ethnic groups there are definite differences with regard to differential birth-rate in the various months of the year. When such groups, which differ also in average mental ability, are put together into one there would result some seasonal differences in the latter factor. Whether the seasonal variation in birth rate would hold true for larger samples of these ethnic groups is still to be determined. For the total population such information is available. Table 2 gives the number of births recorded in New York City, by months,

TABLE 2
NUMBER OF BIRTHS—BY MONTHS
New York City, 1910-1934*

Month of birth	Number	Daily rate
January	265,549	356.9
February	243,221	358.7
March	270,968	364.2
April	254,463	353.4
May	249,586	335.5
June	260,678	362.1
July	265,080	356.3
August	264,450	355.4
September	250,235	347.5
October	255,556	343.5
November	237,225	329.5
December	252,266	339.1
June-November	1,533,224	349.1
December-May	1,536,053	351.3

*The year 1918 was excluded because the data were incomplete for the 25-year period, 1910 through 1934 (the year 1918 is excluded because data for that year were incomplete). This tabulation is based on 3,069,277 birth records in the files of the Department of Health

It will be noted that there was considerable variation and that in the six months, December-May, there was a somewhat larger number of births than in the months, June-November, the difference being 2829 cases The general trend of the daily birth-rate for each of the 12 months during the 25-year period is shown in Figure 1

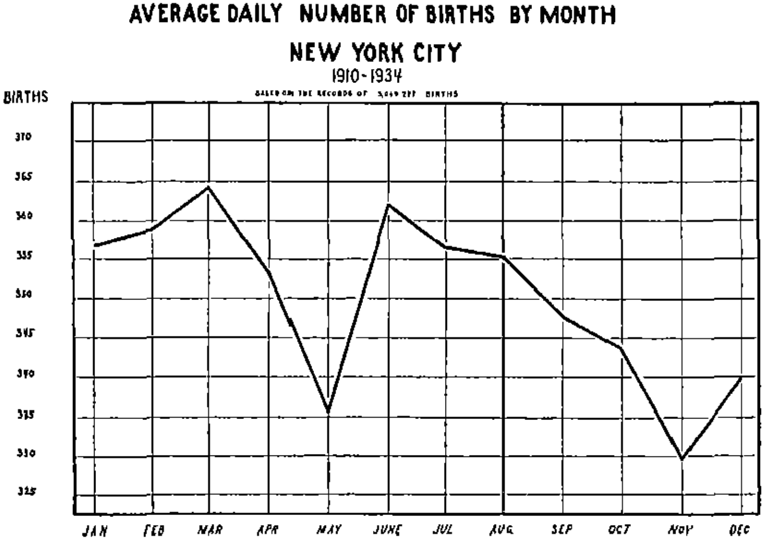


FIGURE 1

Distribution of IQ by Month of Birth

The group of 6353 fifth-grade pupils included in this study was examined with the National Intelligence Test and the Pintner Rapid Survey Test[2] All the testing was done during the same month

---

[2] The authors acknowledge the cooperation of Dr Eugene Nifenecker and Miss May Lazar of the Bureau of Reference, Research, and Statistics, Board of Education, New York City

TABLE 3

IQ AND MONTH OF BIRTH FOR VARIOUS ETHNIC GROUPS

Month of birth	Italian Mean	Italian S.D.	Italian N	Jewish Mean	Jewish S.D.	Jewish N	Negro Mean	Negro S.D.	Negro N	Total Mean	Total S.D.	Total N
January	85.4	17.9	188	103.6	17.3	254	88.2	16.0	103	94.4	19.3	545
February	87.4	17.4	164	104.0	19.8	230	86.2	17.9	114	94.6	20.5	508
March	86.7	18.6	171	104.6	19.1	286	84.8	16.2	123	95.1	20.6	580
April	86.6	16.4	179	105.3	18.3	264	86.6	19.0	107	95.6	20.1	550
May	86.1	18.1	172	103.0	19.0	269	85.9	16.9	123	94.1	20.2	564
June	88.9	18.4	158	104.3	17.5	208	88.3	19.7	86	95.9	19.8	452
July	89.4	17.8	164	105.9	18.1	263	86.8	17.5	128	96.6	19.9	555
August	86.9	17.5	138	104.9	19.0	283	88.5	18.7	116	96.7	20.5	537
September	88.6	17.4	174	105.4	17.1	269	87.1	18.0	115	96.4	19.4	558
October	89.2	19.2	166	104.4	16.6	261	90.2	18.5	114	96.7	19.3	541
November	88.7	18.4	172	105.7	15.7	204	90.5	16.9	98	96.4	18.8	474
December	87.2	16.0	162	103.2	18.5	208	89.6	18.0	119	94.6	19.1	489

AVERAGE IQ BY MONTH OF BIRTH

Table 3 shows the mean IQ, the standard deviation, and the number of cases for each ethnic group and for the total group by month.

The largest and smallest monthly means for each group are as follows.

Population	Highest monthly mean		Lowest monthly mean		Difference	Diff / σ_{diff}
Italian	July	89.4	January	85.4	4.0	2.1
Jewish	July	105.9	May	103.0	2.0	1.8
Negro	November	90.5	March	84.8	5.7	2.5
Total	August-October	96.7	May	94.1	2.6	2.2

The largest monthly means fall in summer and autumn, whereas the smallest means fall in winter and spring. The critical ratio,

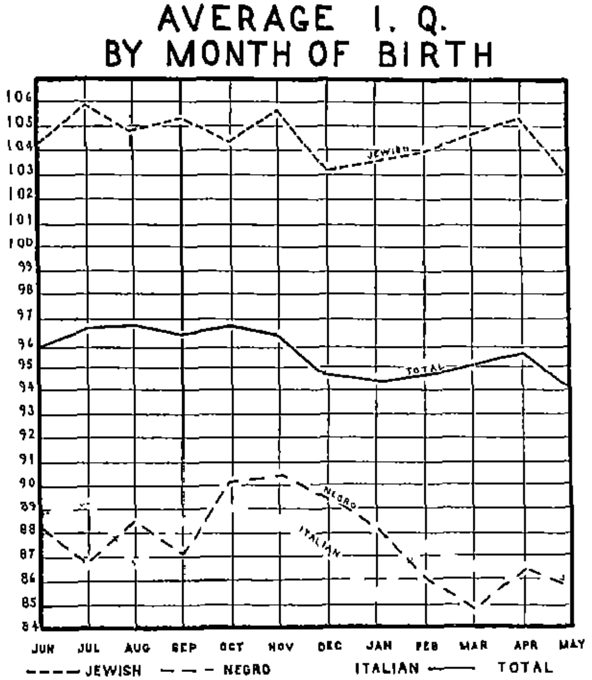


FIGURE 2

the difference in terms of its standard error, is shown in the last column.

The general trend of the averages by month is shown graphically in Figure 2. The most striking is the curve for the total population. This shows consistently higher means for the six months from June to November, inclusive, and consistently lower means for the other six months of the year. The curves for the three ethnic groups are less striking, but the general trend in all of them is rather consistent.

Seasons. If, now, we group the months into seasons, we have the results presented in Table 4. The lowest and highest seasonal means are as follows:

Population	Highest seasonal mean		Lowest seasonal mean		Difference	$\frac{\text{Diff}}{\sigma_{diff}}$
Italian	Autumn	88.8	Spring	86.5	2.3	2.1
Jewish	Autumn	105.1	Winter	103.6	1.5	1.5
Negro	Autumn	89.2	Spring	85.7	3.5	2.6
Total	Autumn	96.5	Winter	94.5	2.0	2.9

We note here that the highest means always occur in autumn, and the lowest means in winter and spring.

The differences in mean IQ are small, ranging from 1.5 to 3.5. However, the difference was consistent in the ethnic groups, and it approached statistical significance in the total population.

Half-Year Groupings. When the months are grouped in half-year intervals, the results are seen in the following figures:

Population	June-November Mean	December-May Mean	Difference	$\frac{\text{Diff.}}{\sigma_{diff}}$
Italian	88.6	86.5	2.1	2.6
Jewish	105.1	104.0	1.1	1.7
Negro	88.5	86.8	1.7	1.8
Total	96.5	94.7	1.8	3.6

The mean IQ for the total population was 96.5 for the period June-November and 94.7 for December-May. The difference is 1.8 and statistically reliable, being 3.6 times its S.D. When the ethnic groups are considered individually, the differences although

TABLE 4

IQ and Season of Birth

Season of birth	Italian Mean	Italian S D	Italian N	Jewish Mean	Jewish S D	Jewish N	Negro Mean	Negro S D	Negro N	Total Mean	Total S D.	Total N
Spring	86.5	17.7	522	104.3	18.8	819	85.7	17.3	353	94.9	20.3	1694
Summer	88.5	17.9	460	105.0	18.3	754	87.8	18.5	330	96.4	20.1	1544
Autumn	88.8	18.3	512	105.1	16.5	734	89.2	17.9	327	96.5	19.2	1573
Winter	86.6	17.2	514	103.6	18.5	692	88.0	17.4	336	94.5	19.6	1542
Winter-Spring	86.5	17.4	1036	104.0	18.7	1511	86.8	17.4	689	94.7	20.0	3236
Summer-Autumn	88.6	18.2	972	105.1	17.4	1488	88.5	18.2	657	96.5	19.7	3117

not statistically significant, are consistently in favor of the June-November period

Comparison of Extremes. The difference between the summer-fall and the winter-spring children expresses itself also in the extremes If we consider the proportions of cases of mental deficiency or feeblemindedness (IQ below 65) and the proportion of cases of very superior intelligence (IQ above 130), we find the following comparison. Among those born in the months of June-November there were fewer cases of mental deficiency than mental superiority, the percentages being 4.6 per cent and 5 9 per cent, respectively Among those born in December-May there were more cases of mental deficiency than of mental superiority, the proportions being 6.1 per cent and 5 3 per cent, respectively[8]

Possible Causes of Variation in IQ We have discussed the possibility that differences in the monthly birth-rates of our ethnic groups may be the cause of the seasonal variation in the average IQ We have seen, however, that this could not have been the sole factor, because the general trend of the variation in IQ held true for practically all the ethnic groups It now remains to consider other possible causative factors The physical stamina of the child may have some connection with his general intelligence Children born during the warm months may have better health records than those born during the cold months Among the most significant indices of child health are neonatal and infant mortality rates

Neonatal mortality (deaths of infants under one month) is seen to be much smaller during the warm months Table 5 shows the number of neonatal deaths in New York City during 1931-1934 The number of deaths for the total and for each ethnic group separately is definitely much smaller during the warm months. There is

---

[8] With reference to the dull and feebleminded we have data collected by Mr. Rosenthal for 8985 cases, with a mean IQ of 65 2 The means vary slightly from month to month The lowest monthly mean falls in February, but the highest falls in May. The lowest mean for three-month groupings is for December-February, but the highest is for March-May The cool months (October to March) have a very slightly lower mean than the warm months (April to September) None of the differences is statistically significant. Out of 8985 cases, 27 more were born in the December-May half-year interval than in the remaining six months These data support the findings of the present study to some extent, particularly in regard to the lower IQ's found among those born in the coldest part of the year

TABLE 5
NEONATAL MORTALITY—BY MONTHS
New York City, 1931-1934.

Month of death	White Total city	Negro Total city	Italian Neighborhoods	Jewish Neighborhoods	Total N.Y C	Ratio
January	1075	123	60	55	1198	1.13
February	858	121	42	43	979	1.01
March	1019	142	54	49	1161	1.09
April	987	131	49	45	1118	1.09
May	1010	99	45	50	1109	1.04
June	872	119	39	36	991	0.96
July	911	99	42	41	1010	0.95
August	934	116	49	41	1050	0.99
September	905	89	43	44	994	0.97
October	888	107	48	47	995	0.94
November	849	109	46	37	958	0.93
December	868	97	43	41	965	0.91
June-November	5359	639	267	246	5998	0.96
December-May	5817	713	293	283	6530	1.04
Total	11,176	1352	560	529	12,528	

much fluctuation from month to month, but the general trend is clear.

At the bottom of the table we note that the number of cases of neonatal mortality in New York City for the summer-autumn months (June-November) was 5998, while in the winter-spring months it was 6530. The ratio of the obtained to the expected was .96 for the summer-autumn months and 1.04 for the winter-spring months. The rates were 28.89 and 30.98 per 1000 births, respectively.

The trend for the ethnic groups was consistent with that for the total population. Among the Italians, the number of cases of neonatal mortality was 293 for the December-May months and 267 for the June-November months; the rates per 1000 births were 24.29 and 26.75, respectively. Similarly, in the Jewish group, the number of cases of neonatal mortality was 283 for the December-May months and 246 for the June-November months, the rates per 1000 births were 25.25 and 22.92, respectively. The rate of neonatal mortality is to some extent an index of the general vitality of the children born. Those born in the winter-spring months who survive have been subjected to greater physical strain than those born in the summer-autumn. This may in turn result in a lower average mental capacity.

Infant Mortality. Applying the same method of analysis as in the case of neonatal mortality, we may study the monthly variation in infant mortality (those dying during the first year of life). Table 6 shows the number of deaths for 1930-1934. Again, we note that infant mortality is definitely lower in the summer-autumn months than in the winter-spring months. During the five-year period, 1930-1934, the number of infants (less than one year old) dying in the warm months was 12,693, while the number dying in the cold months was 16,546. The ratio of the obtained to the expected was .87 for the summer-autumn months and 1.13 for the winter-spring months. The birth-rate in the warm months is slightly lower (a difference of about two cases per day), but the rate of infant mortality is proportionately much smaller. The rate

TABLE 6
INFANT MORTALITY—BY MONTHS
New York City, 1930-1934

Month of death	White Total city	Negro Total city	Italian Neighborhoods	Jewish Neighborhoods	Total N Y C	Ratio
January	2587	349	155	117	2936	1.18
February	2314	318	140	110	2632	1.16
March	2658	355	164	141	3013	1.21
April	2613	322	163	124	2935	1.22
May	2481	307	151	123	2788	1.12
June	1900	269	100	76	2169	0.90
July	1851	226	95	80	2077	0.84
August	1900	253	121	84	2153	0.87
September	1846	200	109	86	2046	0.85
October	1891	244	110	87	2135	0.86
November	1865	248	115	79	2113	0.88
December	2039	203	153	107	2242	0.90
June-November	11,253	1440	650	492	12,693	0.87
December-May	14,692	1054	926	722	16,546	1.13
Total	25,945	3294	1576	1214	29,239	

of infant mortality in the warm months was 47.28 per 1000 live births and in the cold months 61.12 per 1000 live births

This relationship is true also for the two ethnic groups. Among the Italians, the number of infants dying in the warm months was 650 as compared with 926 in the cold months The rates per 1000 births were 43.91 and 62.74, respectively In the Jewish group, the number of infants dying in warm months was 492 as compared with 722 in the cold months. The rates per 1000 births were 34.55 and 48.90, respectively.

Morbidity No comprehensive statistics of illness in infancy by month of year have been found We have, therefore, taken the records of mortality due to congenital and early infancy diseases as another measure of the general health of infants born during different months of the year The months showing the greatest number of deaths from these causes are taken to be those in which the general morbidity was highest.[1] This is shown in Table 7 During the six-year period, 1929-1934, the number of children dying from con-

TABLE 7
DEATHS DUE TO CONGENITAL AND EARLY INFANCY DISEASES—BY MONTHS
New York City, 1929-1934

Month of death	White	Negro	Total	Ratio
January	1608	193	1801	1.11
February	1363	165	1528	1.03
March	1601	196	1797	1.11
April	1466	206	1672	1.07
May	1561	153	1714	1.06
June	1313	166	1479	0.94
July	1404	151	1555	0.96
August	1390	173	1563	0.96
September	1375	129	1504	0.96
October	1365	158	1523	0.94
November	1284	155	1439	0.92
December	1377	141	1518	0.94
June-November	8131	932	9063	0.95
December-May	8976	1054	10,030	1.05
Total	17,107	1986	19,093	

---

[1] This holds true also for certain children's diseases such as diphtheria and scarlet fever These were more prevalent in the winter-spring months than in the summer-autumn months See Annual Report of Department of Health, City of New York, 1928

genital and early infancy diseases during the summer-autumn months was 9063 as compared with 10,030 for the winter-spring months, a difference of 967 cases The ratio of the obtained to the expected is 95 for the June-November months and 1.05 for the December-May months For the ten-year period, 1925-1934, for New York City the number of deaths (due to congenital and early infancy diseases) for the June-November months was 15,738 and for the December-May months 17,432, a difference of 1694 cases

We have also tabulated the deaths due to puerperal diseases as an index of the health of the mother. Health of the mother at the time of birth probably affects the health of the offspring. Table 8 shows the number of deaths for the years 1931 to 1934 The total number of deaths per month is much less during the warm months than during the cold months During 1931-1934, the number of deaths in the June-November interval was 1090 and in the December-May interval, 1409 The ratio of the obtained to the expected, in the last column, shows a larger figure for every month during the December-May interval as compared with the June-November interval The mothers of children born during the June-November interval are apparently less weighted with illness than the mothers of children born during the other half of the year.

TABLE 8
DEATHS DUE TO PUERPERAL DISEASES—BY MONTHS
New York City, 1931-1934

Month of death	Number	Ratio
January	238	1 12
February	233	1 21
March	261	1 23
April	237	1 15
May	229	1 08
June	191	0.93
July	159	0.75
August	206	0.97
September	198	0.96
October	142	0 67
November	194	0.95
December	211	0 99
June-November	1090	0'87
December-May	1409	1 13
Total	2499	

Comparison with Previous Findings As stated in the early part of this report, Blonsky in 1929 studied 453 backward children and found the highest mean IQ among children born in spring Pintner in 1931 studied 4925 cases and found the lowest IQ among those born in winter. Pintner and Forlano in 1933 made a more extensive study with 17,502 cases divided into children of low, median, and high social status They found the lowest mean to occur in the winter months (January to March) for all three groups of children The highest mean occurred in the spring months (April to June).

The findings of the last two studies are, to a certain extent, similar to our present findings Children born in the coldest months of the year were found, on the average, to have slightly lower IQ's In the present study we find that the drop in IQ begins in December and continues through to May. The Pintner and Forlano study showed that the drop also begins in December but continues only until March. With references to the spring months there is, therefore, no agreement The one consistent finding is the lowered IQ in the months from December to March This, in general, is true, whether we divide our cases into different ethnic groups as we have done in this study or into different social levels as in the previous study Although the differences in IQ's are small, they seem to be consistent and, in some of the comparisons, statistically reliable

SUMMARY

1. Children born in the summer-autumn months (June through November) show a higher average IQ than those born in the spring-winter months (December through May) This difference though small (1 8) is statistically significant, 3 6 times its standard error Each of the six months of the former group has a higher average than each of the months in the latter group The lowest of one group is higher than the highest of the other group This general difference in favor of the summer-autumn months held true in each of the ethnic groups studied · Italian, Jewish, and Negro

2 When the data are considered by individual months (or periods smaller than half-year intervals) and by ethnic groups, the results are not so consistent as for the total population. This indicates the need of considering the ethnic factor in studies of this kind. The comparison between the warm and the cold months, however,

do show a consistent difference in favor of the warm months, June-November, for all the ethnic groups under consideration.

3. Vital statistics favor the warm months (June-November). These months show a lower rate of infant mortality, neonatal mortality, deaths due to congenital and early infancy diseases, and deaths due to puerperal diseases.

Before general conclusions are drawn from this investigation, it should be remembered that our population consisted of children all of whom were enrolled in Grade V of the elementary schools of New York City. This homogeneity, though advantageous from the point of view of reducing the number of varying factors, might in itself be a selective factor causing seasonal variations. Thus, the children within a given school grade, those whose birthdays fall in one part of the year, might as a group be slightly older than those born in another part of the year. The older children are usually duller than the younger ones within the same grade. This might result in an apparent relationship between the season of birth and average intelligence.

In the "population" described in this study, there seems to be evidence that those born in the warm months (June-November) are, on the average, brighter than those born in the cold months (December-May). This general difference has been noted in each of three ethnic groups: Italian, Jewish, and Negro children. The same difference in favor of the warm months was found to hold true in several aspects of infant vitality, as revealed in health statistics.

The ethnic character of the population seems to be an important factor in studies of seasonal variation. Many of the contradictory results in the realm of physiological and anthropometric characteristics of infants born in different seasons are probably due to differences in the ethnic character of the population on which those studies were based.

REFERENCES

1. Abel, H. Weight at birth and vitamins according to publications from years 1923-1925. *Arch. f. Kinderh.*, 1926, 78, 187.
2. Andersen, H. Sermo de pondere et longitudine infantum recens natorum. *Nordiskt Med. Ark., N. F.*, 1899, 10, 24.
3. Bakwin, H., & Bakwin, R. M. Body build in infants. V. Anthrometry in the new-born. *Human Biol.*, 1934, 6, 612-626.

4 ——— Seasonal variation in the weight loss of new-born *Amer. J Obstet & Gynec*, 1929, **18**, 863.

5 Blonsky, P P Fruh- und Spatjahrkinder *Jahrb f Kinderheilkunde,* 1929, **124**, 115-120

6 Brenton, H Climate and race as factors influencing the weight of the new-born *Amer J Phys Anthrop*, 1922, **5**, 237-249

7 Fairgrieve, McC Birthdays in relation to intelligence *Nature,* 1922, **109**, 218.

8 Hansen, J H Researches upon the weight of new-born children *Meddel f. Anthrop Kom* Copenhagen, 1913 Pp 109

9 Hellmuth, K. Sind Jahreszeitliche Schwankungen der Geburtsgewichte Statistisch Nachweissbar? *Zsch f Geburtsch u Gynaek*, 1928, **93**, 147

10 Hermann, C The seasonal incidence of some of the diseases of childhood

11. Kassel, C Birth month of genius Chicago Open Court, 1929

12 Maller, J B Vital indices and their relation to psychological social factors *Human Biol*, 1933, **5**, 94-121

13 Pintner, R Intelligence and month of birth *J Appl Psychol*, 1931, **15**, 149-154

14 Pintner, R, & Forlano, G The influence of month of birth on intelligence quotients *J Educ Psychol*, 1933, **24**, 561-584

15 Sanders, B S Environment and growth Baltimore, Md Warwick & York, 1934 Pp 375.

16 Smith, S Influence of illness during the first two years on infant development *J Genet Psychol.,* 1931, **39**, 284-287

17 Woodbury, R M Causal factors in infant mortality. *Children's Bureau Publ*, No 142, U S Dept of Labor, Washington, 1925 Pp 245

Teachers College
Columbia University
New York City

A METHOD FOR THE QUANTITATIVE ANALYSIS OF HUMAN GAIT*[1]

Training School, Vineland, New Jersey

GEORGE KREEZER AND A. DOUGLAS GLANVILLE

A. INTRODUCTION

In the present paper, a method is described for the quantitative analysis of human gait. The method has been devised in an attempt to provide ways for the analysis of complex motor functions in general. It was felt that the possibilities in this direction might be better envisaged if concrete data were first obtained in connection with a particular case of complex motor activity. Walking was selected as an example of an important motor activity which lent itself to analysis by means of the type of method which we wished to examine.

The central idea behind the method of analysis proposed is that all complex motor activity consists essentially of a temporal organization of movements of which the components are the positions and movements at individual joints. The task of describing such motor activities becomes, then, that of specifying the particular positions and movements which occur at the various joints involved, and of indicating the manner in which these component movements are organized in the total act. A description of this sort should tell us of the precise way in which the "bodily machinery" operates during motor activity, and should provide the data on the basis of which the underlying neurological mechanisms may be investigated.

In the case of walking, then, our problem has been to find ways of

---

*Received in the Editorial Office on July 13, 1936.

[1] We should like to express our appreciation to Dr. Edgar A. Doll, Director of the laboratory, who, on the basis of his previous investigations of the birth-injured, encouraged us to develop methods for the quantitative analysis of human motor functions. A condensed version of the present paper was read at a meeting of the New York Branch of the American Psychological Association, April, 1936.

measuring the principal component movements and positions at individual joints, and to determine the manner in which they are organized or intercoordinated in the course of the walking sequence The provision of ways for determining these facts is the principal aim of the present method[2] It is also important, however, in the analysis of gait to take into account the relation of the individual to gravitational forces Consequently, properties concerned with the nature of the base of support and its relation to angular movements at various joints have also been considered.

B Techniques Used in Analysis of Gait

The principal steps involved in the present method for the analysis of gait are as follows: (1) Motion pictures are taken of the subject walking under certain standard conditions (2) The individual motion-picture frames are projected one at a time, and measurements made of the angular positions at hip-, knee-, and ankle-joints On the basis of these measurements, graphs are constructed showing the changes taking place concurrently at hip, knee, and ankle-joints as a function of time, as the subject walks (Figure 2) (3) On the basis of these graphs and of further measurements upon the individual frames, a large number of functionally important properties of gait are determined in quantitative terms (4) As a supplement to these quantitative determinations, an examination of the subject's gait is also made on the basis of the motion pictures projected at the standard rate, and an estimate made of the nature of certain characteristics of gait observable under these conditions These estimates provide what may be regarded as a qualitative description of the walking sequence which may be compared with the results of the quantitative procedures referred to above They also provide information concerning certain characteristics of gait for which we have not yet provided quantitative methods of determination We may now examine in more detail the various procedures used in connection with the steps just mentioned.

1. *Conditions Under Which Motion Pictures Are Taken.* The subject is instructed to walk at a moderate rate upon a thin mat,

---

[2] For a preliminary description of this method and its place in a more general program for the investigation of the physiological basis of human motor activities, see Kreezer (10)

marked off with reference lines, which is laid upon the floor. Photographs are taken of profile, front, and back views of the subject's gait, that is to say, views in which the subject walks back and forth at right angles to the optical axis of the camera, and views with him walking directly toward and away from the camera It is desirable, during the taking of the picture, that the subject wear some type of costume that will not obscure the body contours. In the taking of the profile pictures, the camera was located at a distance of about 30 ft from the subject Its height was adjusted to correspond about to the level of the hip joint In the investigation of adult subjects, it was kept constant at a level of 3 ft 6 in. In the taking of the profile pictures, the subject walks over a distance of about 15 ft., this distance subtends an angle of about 30° at the camera In the taking of the front and back views, the subject walks between points at distances of about 24 ft and 40 ft. from the camera

The mat upon which the subject walks is marked with reference lines to indicate his location at any given moment. The mat is about 18 ft by 4 ft in size, dull black in color, and marked with white lines which cross each other at right angles Lines running parallel to the direction of progression are referred to as longitudinal lines, and lines at right angles to these, as transverse lines The distance between longitudinal lines is 25 cm and between transverse lines, 50 cm Behind the subject a dark curtain is hung to provide a background Against it hang a number of white-tape plumb-lines, placed at regular intervals to indicate in the pictures the exact direction of the vertical.

At one side of the field of view is placed a clock-dial, marked off in 16 equal divisions. A pointer, driven by a Telechron synchronous motor, revolves around this dial once a second This clock is used to provide a check on the camera speed, and is a method for numbering the successive motion-picture frames A standard camera speed of 16 frames per sec was generally used, a speed of 32 frames per sec may be used, however, when shorter time intervals between successive pictures are desired The nature of the pictures obtained under the standard conditions is illustrated in Figure 1

As indicated above, a walking distance of 15 ft was used in the

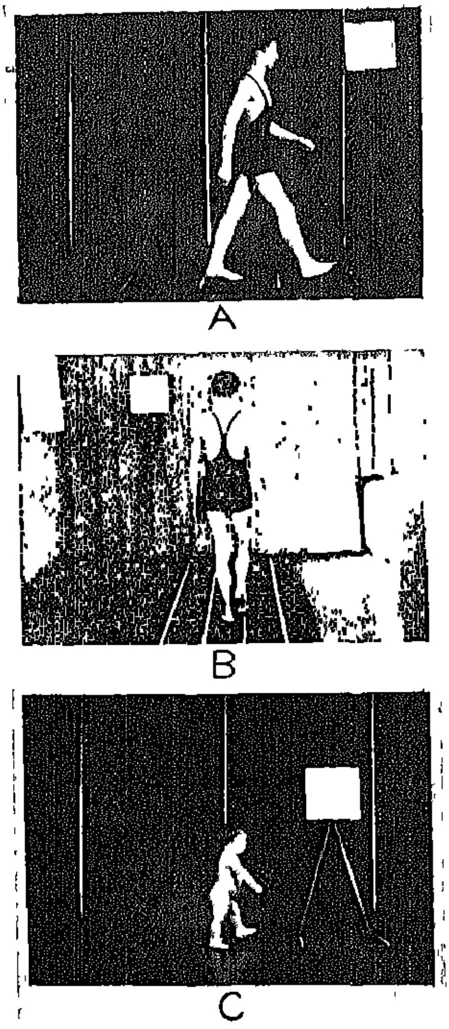


FIGURE 1

SAMPLE OF MOTION-PICTURE FRAMES USED IN THE ANALYSIS OF GAIT

(*A*) Profile view used in measurements of joint-angle (adult subject); (*B*) back view used in measurements of width of base, (*C*) profile view of child shortly after learning to walk without support taking of profile pictures. This range was found sufficient for providing the data considered necessary in our initial, exploratory studies of gait. For certain studies of gait, by means of the method described, it might be desirable to provide for pictures of the subject walking over a greater distance.

2. *Measurements of Joint-angle and Resultant Graphs.* The successive motion-picture frames obtained of the gait of the subject in profile are projected, and measurements made of the angular positions present at hip-, knee-, and ankle-joints, respectively, of each of the two legs. In the establishing of normal adult standards, the pictures used were those in which the right side of the subject is toward the camera. The particular angles measured at each of the joints is indicated in the diagrams of the various joints shown in Figure 2. Measurements are made by means of a large celluloid protractor or goniometer designed for the purpose. A diagram

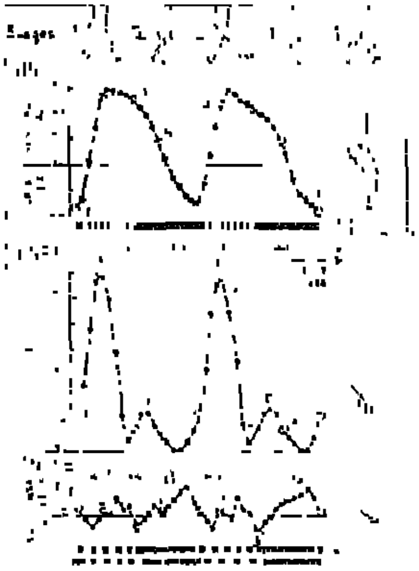


FIGURE 2

JOINT MOVEMENTS IN WALKING

Normal adult—right leg (explanation in text)

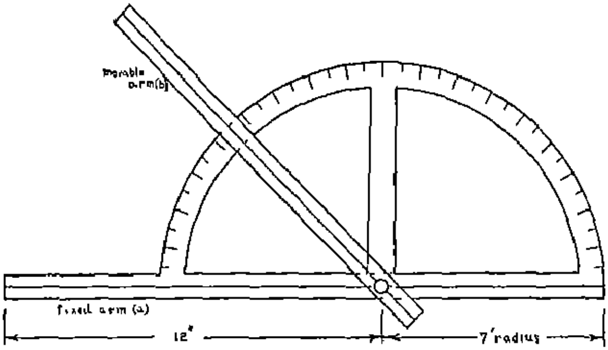


FIGURE 3
Diagram of Protractor Used in Measurements of Angular Position at Various Joints
Arc is graduated in half-degree units

of this protractor is shown in Figure 3[3] Measurements are facilitated, if the screen used for projection of the pictures is rigid and stable, as provided, for example, by a large sheet of cardboard placed flat against a solid wall To permit accurate use of the protractor, the projected pictures should not be too small A scale of projection in which the height of the individual was represented by a 10-in extent on the screen was found to be satisfactory and was adopted as a standard condition in our investigations.

The procedure found to be most accurate and convenient in the carrying out of the measurements may be illustrated by the steps followed in making measurements at the knee-joint The protractor is placed flat against the screen and adjusted until the fixed arm (*a*) of the protractor coincides approximately with the central axis of the thigh, and is parallel to it. The movable arm (*b*) of the protractor is then moved until it is parallel to the axis of the leg. If it does not coincide approximately with the central axis, it is either because the fixed arm is displaced to the side of the axis of the thigh or because the end of the fixed arm of the protractor is too high or too low, so that the pivot does not lie approximately over the fulcrum of the joint (Figure 4C). Adjustment of the position of the fixed protractor arm may be made accordingly, and the movable arm adjusted again. It is apparent, however, from Figure 4,

---

[3] A protractor of the sort described was made for us by Theodore Alteneder and Sons, Philadelphia.

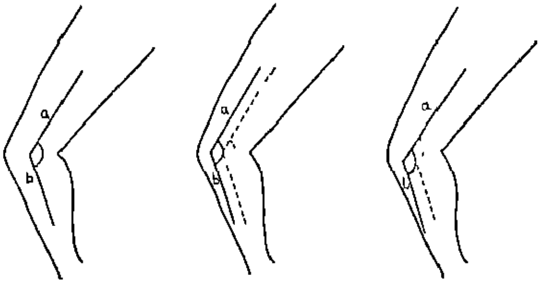


FIGURE 4

DIAGRAM OF USE OF PROTRACTOR IN THE MEASUREMENT OF JOINT-ANGLE

Diagram illustrates the fact that the essential requirement for a correct measurement is that protractor arms be parallel to the axes of the body-members adjacent to the joint. (Further explanation in text.)

that it is not necessary, in obtaining a measure of a given joint-angle, for the pivot of the protractor to lie over the fulcrum of the joint, or for the center-lines of the protractor arms to coincide with the estimated axis of the thigh and the leg of the subject. The accuracy of measurement tends to be greater, however, if the protractor arms coincide approximately with the central axis of the members adjacent to the joint. The use of a transparent material in the construction of the protractor and of relatively thin, black lines engraved down the center of the arms aids in the making of accurate measurements, since they reduce to a minimum the distortion and blocking out of the contours of the body-members. These contours seem to provide the most valuable perceptive cues in the accurate placement of the protractor arms. After some practice, a considerable degree of accuracy may be attained in the measurement of joint-angles, as indicated by the figures on the probable error of a single measure reported below.

On the basis of measurements made of successive frames, graphs such as those shown in Figure 2 may be constructed. These graphs exhibit the changes taking place concurrently at hip-, knee-, and ankle-joints as a function of time for the right leg. The use of a common time axis provides a bird's-eye view of the temporal organization of the movements at the various joints and aids in the determination of the interrelations between these movements in the carrying out of the complex activity We have not come across any previous report in the literature of a method for providing such graphic records of angular changes at the joints during walking

It is convenient to mark out certain points and intervals on the graphs on the basis of their functional significance in the act of walking We may consider, first, certain obvious features of the walking sequence As an individual walks, the two legs alternate in their support of the body. The interval during which a given leg is in contact with the ground may be called the *ground stroke;* the interval during which it moves forward through the air, the *recovery stroke* The moment the left foot first strikes the ground, at the beginning of its ground stroke, is labeled *Stage I,* the corresponding moment at the beginning of the ground stroke of the right leg, Stage III Stages II and IV refer to intermediate moments, Stage II occurring during the ground stroke of the left leg, and Stage IV during the ground stroke of the right leg These intermediate stages are a little difficult to describe, although readily recognized in the pictures. The essential feature is that the body is supported on one leg in an apparently stable position relative to gravity A vertical line drawn from the trunk to the ground would pass through the base of support, and the various parts of the body seem to be symmetrically distributed about this vertical axis The particular frames at which the various stages occur is determined by moving along by hand, at a moderate speed, the successive frames in the projector. A complete cycle of change in the walking sequence consists of an interval lying between two successive repetitions of the same stage.

Underneath the set of curves showing joint-angle as a function of time, two horizontal lines are drawn to indicate the time during which the heel and toe, respectively, are in contact with the ground. The horizontal line drawn under the curve for the hip-joint shows the time during which any part of the foot is in contact with the ground It may, therefore, be regarded as a resultant of the heel and toe records A solid line indicates contact with the ground, a broken line indicates lack of contact The "foot-contact" line indicates the duration of the ground stroke, since the ground stroke, by definition, represents the interval during which any part of the weight-bearing leg is in contact with the ground.

A set of graphs similar to those show in Figure 2 is constructed to represent movements occurring at the joints of the left leg. The form of the curves is similar, there is, however, a displacement in phase—the angular positions occurring in the right leg at Stage I, not occurring in the left leg until Stage III, one-half cycle later, has been reached

3. *Procedures for Measuring Various Properties of Gait.* We may now consider the various properties of gait which can be determined in a given subject, and the specific procedures used for the purpose

a Extent of joint-movement during a single cycle The extent of joint movement during a cycle may be specified for each joint in terms of (*a*) the maximal angular position reached, (*b*) the minimal angular position reached, and (*c*) the maximal range of angular motion occurring during a cycle beginning at Stage I These characteristics are determined by reference to the graphs to identify the frames in which the maximal and minimal angular positions occur and by reference to the data tables for the exact values By maximal position is meant the highest point reached in the curve for a joint during a single cycle, and by minimal position, the lowest point reached. The terms are not synonymous with the mathematical use of the terms *maximum* and *minimum,* which stand for places where the slope of the curve is reversed in passing through a point of zero slope It is apparent that a mathematical maximum, as in Frame 12 of the curve for the knee-joint (Figure 2), is not necessarily the highest point reached during a complete cycle The characteristics we are concerned with here are the highest maximum and the lowest minimum The maximal *range* of angular movement refers to the difference between the maximal and minimal angular positions. In determining these three characteristics for our table of norms, the initial cycle in the graph, lying between successive Stage I points, is used

b Form of the curves representing angular changes at the joints Ideally, the form of the curves shown in Figure 2, would be specified by mathematical equations We have not tried, however, to derive equations for describing these curves, but have limited ourselves to a description in empirical terms An interval of the curve covering a single cycle is first labeled in accordance with the number of waves appearing throughout a cycle By a wave is meant the part of the curve lying between two successive troughs or mathematical minima Waves which are less in amplitude than 4°, measured from trough to peak, are not considered, in view of the determinations made that the probable error of a single measure is about 1° for measurements at hip-, knee-, and ankle joints

Examination of the curves of a group of normal adults shows, as is also apparent in the curves of Figure 2, that at the hip-joint only a single wave occurs during a cycle, at the knee-joint two waves, and at the ankle-joint two fundamental waves, though there may be smaller waves superimposed upon the larger waves. In the case of the curve for the knee-joint, the initial wave, beginning about half way between Stages IV and I is called the *"A"* wave, and the succeeding small wave beginning just before Stage III is called the *"B" wave* Similarly in the case of the ankle, the wave beginning shortly after Stage I is called the *"A"* wave, and the wave beginning at Stage III, the *"B"* wave.

The following characteristics of the individual waves may now be measured. These characteristics have been selected as providing an empirical description of the form of curves of the sort found among normal subjects The characteristics examined are amplitude of the rising limb of the wave (a), amplitude of the succeeding falling limb (b), the difference between amplitudes of the rising and falling limbs (a-b), ratio of rising and falling limbs $\left(\frac{a}{b}\right)$, the phase of the maximum point of the wave $\left(\frac{c}{P}\right)$ and the fraction of the total period occupied by the total wave $\left(\frac{d}{P}\right)$. The meaning of these terms is illustrated in Figure 5. The rising limb of the wave represents the interval between the peak of the wave and the preceding trough. The amplitude may be readily determined by subtracting the angular position representing the trough from that representing the peak In the case of the hip, the rising limb of any wave represents an interval of ventral flexion; in the case of either of the "knee-waves," an interval of flexion; and in the case of the "ankle-waves," an interval of dorsi-flexion

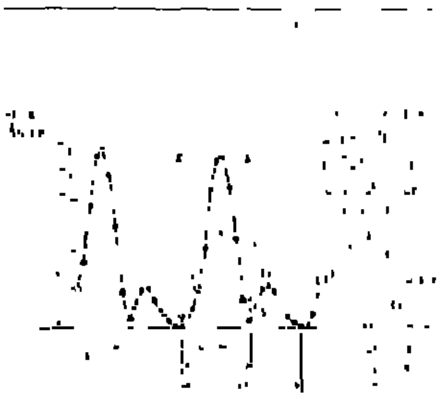


FIGURE 5

Diagram Showing the Characteristics Which Are Measured upon a Given Wave in the Determination of the Form of the Curves of Angular Movements at the Joints

The curve shown is of movement at the knee-joint.

Similarly, the falling limb of a wave includes the interval between the peak and the succeeding trough, and represents an interval of dorsi-flexion for the "hip-waves," extension for the "knee-waves," and plantar flexion for the "ankle-waves."

The *phase* at any point of the wave is defined as the fractional part of the cycle that has passed by the time that point on the wave has been reached—the trough at which the "*A*" wave begins being regarded here as the beginning of the cycle. The phase of the maximum point of the wave may, thus, readily be determined by dividing the time required in the cycle for the maximal point on the wave to be reached by the *period* of the curve. The *period* is defined as the duration of one complete cycle, in accordance with the conventional use of this term in describing wave properties. The use of the term *phase* here is also analogous to its mathematical use, as the term is used, for example, in the description of sound waves and alternating electrical currents and potentials.

c. *Angle of foot axis.* By the angle of the foot is meant the angle lying between the long axis of the foot and the direction of progression of the individual, at the times when the foot is in contact with the floor, as illustrated in Figure 6*A*

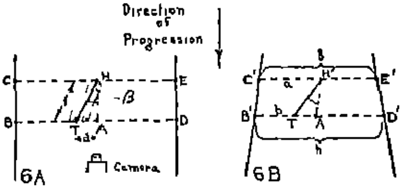


FIGURE 6

Diagrams to Illustrate the Method Used in Measurement of the Direction of the Foot Axis Relative to the Direction of Progression ("Angle of Foot Axis")

A number of earlier workers have investigated this characteristic of gait by obtaining records of the foot prints made during the course of walking. Vierordt (21) attached styluses at different parts of the shoe and thus obtained records of the places where different parts of the foot made contact with the floor Dougan (5) obtained records with hob-nail shoes and carbon paper. Shirley (19), in a recent investigation of motor development in children, covered the bottoms of the feet with an oily solution and obtained records of the footprints upon paper A similar procedure was used by Burnside (2)

It is possible, however, to dispense with such procedures and to determine the angle of "toeing-in" and "toeing-out" during walking on the basis of measurement made from the moving-picture records, coupled with simple trigonometric calculations. The measurements necessary are made upon the moving-picture frames obtained with the subject walking directly toward the camera The method of calculation is illustrated in Figure 6

Line *HT* or *f* represents the length of the axis of the foot *H* represents the heel, and *T* the end of the second toe Line *HA* is a line paralled to the direction of progression and line *TA* or *d* represents the length of the perpendicular dropped to the line *HA*. The angle β may readily be determined if the lengths of lines *f* and *d* are known, since $\beta = \text{arc sin} \frac{d}{f}$ If *d* and *f* are known, the value of the angle may be determined from a table of trigonometric sines

The length f may be most easily determined by direct measurements of the length of the subject's foot. The distance d can be determined from the motion-picture records. In doing so, however, it is necessary to take into account the scale of reduction of extents in the projected pictures relative to the true extents, and the fact that the scale of reduction varies with lines lying at different distances from the camera.

These facts are represented in Figures 6*A* and 6*B*. Figure 6*A* represents the true angles and extents associated with the position of the foot; Figure 6*B*, the appearance of Figure 6*A* as seen in perspective in the projected motion-picture frame. Corresponding points on the two figures are labeled with the same letters, those in Figure 6*B* being primed. Lines *CB* and *ED* represent longitudinal reference lines marked on the mat over which the subject walks. Their distance apart is 25 cm.

Our problem is to determine the length of *TA* or d on the basis of measurements made upon the projected picture (6*B*). A method for doing so is given by the following steps:

$$d = CH - BT \tag{1}$$

$$CH = \frac{a}{g}\,(25) \tag{2}$$

$$BT = \frac{b}{h}\,(25). \tag{3}$$

The last two steps consist simply in determining the true lengths corresponding to the lengths a and b measured on the projected pictures. (The true distance corresponding to extents g and h is known to be 25 cm.)

$$d = 25\left(\frac{a}{g} - \frac{b}{h}\right) \tag{4}$$

$$\beta = \text{arc sin}\ \frac{d}{f} \tag{5}$$

$$\beta = \text{arc sin}\ \frac{25}{f}\left(\frac{a}{g} - \frac{b}{h}\right) \tag{6}$$

In determining the angle of the foot axis at each successive step, it is necessary to measure four distances on the screen and the length of the foot.

d. Width of the base of support. By the width of the base of support is meant the transverse distant apart of the two feet in successive contacts with the floor, measured from the location of the centers of the heels. This characteristic is illustrated in Figure 7*A*, which represents the position of the two feet on the ground as seen by an observer looking down on the floor from above. The width of the base of support is represented by the distance *w* in the figure. This characteristic of gait, like the angle of the foot axis, has been determined by a number of previous investigators on the basis of records taken of the footprints made during the act of walking. Reports are given, for example, by Vierordt (20) Burnside (2), Shirley (19), and McGraw (11) concerning the magnitude of this characteristic under different conditions. The width of the walking base may, however, also be determined on the basis of the motion-picture records alone.

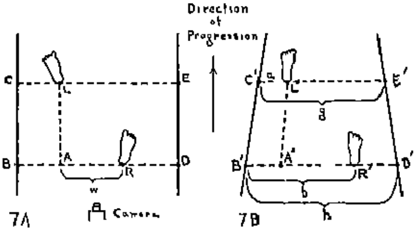


FIGURE 7
DIAGRAMS TO ILLUSTRATE THE METHOD USED IN MEASUREMENT OF THE WIDTH OF THE BASE OF SUPPORT

The pictures used for this purpose are those taken with the subject directly in front of the camera and walking away from it. Measurements are based on the points where the heels make successive contacts with the floor. The method for determining the magnitude of *w* in Figure 7*A* is similar to that used for determining the distance *d* in Figure 6, as described in the previous section. The method is illustrated in Figures 7*A* and 7*B*, and is designed to take into account the difference in the scale of reduction in the projected picture, of transverse lines at different distances from the camera.

Figure *7A* represents the locations of the two feet as they would be seen by an observer looking down at the floor from above, Figure *7B*, the appearance of the same view as projected on the screen from a motion-picture record.

Corresponding points in the two figures are labeled with the same letters, those in Figure 7B being primed Lines CB ($C'B'$) and ED ($E'D'$) represent *lines on the mat* running parallel to the direction of progression They lie at a distance of 25 cm apart Points L (L') and R (R') indicate the points of contact of the centers of the left and right heels, respectively, with the floor.

Our problem is to determine the distance w (Figure *7A*) on the basis of measurements made of extents represented in Figure *7B* A method for doing so is indicated by formula (5) below It is derived as follows.

$$w = BR - CL \tag{1}$$

$$BR = \frac{b}{h}(25) \tag{2}$$

$$CL = \frac{a}{g}(25). \tag{3}$$

$$BR - CL = \frac{25b}{h} - \frac{25a}{g} \tag{4}$$

$$w = 25\left(\frac{b}{h} - \frac{a}{g}\right). \tag{5}$$

All measurements are in cm. It thus becomes necessary in the pictures to measure only distances b, h, a, and g. Substitution in formula (5) gives the corresponding value for the width of the base for that particular step In making these measurements, it is obvious that the heels of the right and left feet will not ordinarily be in contact with the floor at the same time Consequently, the frame showing the right heel in contact is first projected and appropriate measurements made, and then the film is moved along until the frame appears in which the left heel makes contact, and so on for a number of successive steps.

e Distances covered by successive steps The characteristics included under this heading are length of step and what we call the wave-length, in analogy to measurement of wave-length in waves propagated through space, as in the case of lightwaves and soundwaves. The profile pictures showing the subject walking at right angles to the camera are used in the determination of the length of step. The method of measurement employed is exactly like that used in the determination of the width of the base as already described Length of step is defined as the longitudinal distance between the toe of the rear foot and the heel of the forward foot

By wave-length is meant the *space* traversed between the Stage I points of two successive cycles It thus represents the space traversed between the time the left heel first strikes the floor in two successive cycles. It may be determined by taking the sum of the lengths of the steps ending at Stages III and I and the lengths of the two feet

f. Period The period represents the time required for a complete cycle and is determined by counting the number of motion-picture frames occurring between two corresponding points of two successive cycles, e g, between two successive points representing Stage I. Each frame interval corresponds to 1/16 sec. A check on the speed of the camera is provided by the record made on each frame of the position of the clock-pointer on the dial

g. Fraction of a period for various stage intervals The durations of the Stage intervals I to II, II to III to IV, IV to I, and the related intervals I to III, and III to I are readily determined from the graphs of the angular changes at the joints as a function of time The ratio of the time required for each of these intervals, to the total *period* of a cycle, indicates the *relative* time occupied by the various intervals of activities making up an entire cycle of change.

h Velocity. The velocity of the gait may be determined by dividing the *wave-length* or distance traversed during a single cycle by the *period* or duration of the cycle in seconds. A knowledge of the velocity of gait is of interest principally in relation to the nature of the gait mechanisms associated with a particular velocity in a given type of subject.

i. The position of the trunk relative to the vertical axis. This measurement is made by determining the angle made by the axis of the trunk with a line perpendicular to the floor. The direction of the vertical is represented in the pictures by the white plumb-lines which lie against the black background This measurement might obviously be made in each successive frame, just as was done for the angular position at various joints We have limited ourselves to determining the position of the trunk relative to the vertical at each of the stage positions

4 *A Checking Scheme for Qualitative Description of Gait* In our investigations of the mechanism of gait, a procedure for description of various features of gait on the basis of "qualitative estimates," has been used to supplement the quantitative procedures described above The schema used for this purpose permits a systematic, though rough, appraisal of a large number of the characteristics of an individual's gait It may, therefore, be used in certain investigations as a preliminary to more precise quantitative determinations, or to provide a means for a rapid description of an individual's gait when time is not available for the use of the quantitative procedures

It will be noted that the schema given below includes certain characteristics for which quantitative methods of determination have been described above, and also others for which we have not yet worked out quantitative methods. For many of them, quantitative methods of determination similar to those already described might readily be devised.

A checking procedure has been found to be most convenient in the use of the schema. Each characteristic of gait listed is examined in turn and the particular variant of that characteristic found to be present is checked. Where the different variants proposed are not adequate for the description of an individual's gait, additional notes may be added for the purpose

It will be noted that, in the case of many of the characteristics listed, the variants mentioned presuppose a knowledge on the part of the observer of the characteristics of normal (adult) gait The schema cannot, therefore, be used with much assurance of accuracy until quantitative determinations have first been made of the limits of variation in the gait of normal adults, these results used in the definition of the various categories, and the observer trained in their use Prior to the acquisition of such data, the observer must depend on his impressions of the characteristics of normal gait as obtained from gross observation, and he cannot attribute any great reliability to his reports upon individuals examined on this basis It is apparent that the observations necessary for the use of the schema might be made either of motion pictures taken of gait or directly upon an individual walking

The list of characteristics proposed for examination follows After the name of each characteristic, a list of variations in this characteristic that may be present in the gait of a particular individual is given The first-named variant is that presumed to be present in normal adult gait

- I *Relation of Feet to Ground*
  - *A* Width of base of support, normal width, unusually wide, unusually narrow
  - *B* Direction of toes (angle of foot axis) point straight ahead, point in, point out
  - *C* Length of step normal length, unusually long, unusually short
- II *Movements of Legs*
  - *A*. *Recovery stroke*
    - 1 "Take off" from toe, flat-footed.
    - 2 Clearance of ground by foot normal clearance, foot drags, excessive clearance (as when subject steps too high)
    - 3 Manner of clearing
      - *a* Movements at hip flexes with leg in longitudinal plane, flexes with leg in non-longitudinal plane, circumducts thigh.
      - *b* Movements at knee knee flexed, knee straight
      - *c*. Movements at ankle foot in longitudinal plane, foot not in longitudinal plane
  - *B* *Ground stroke*
    - 1. Forward contact heel first, flat-footed, toe first
    - 2. Leg form normally straight, excessively flexed, excessively stiff
- III *Movements of Arms*
  - *A* Range of swing: normal range, excessive range, deficient range (or absent), arms held outstretched
  - *B* Direction of swing slightly diagonally across body, abnormal direction (specify nature)
  - *C* Timing with legs synchronized with opposite leg, not normally synchonized (specify nature)

- *D.* Tension in arms: normally relaxed, stiff.
- IV. *Position of Head*
  - *A.* Profile view: normally erect, ventrally flexed, dorsally flexed.
  - *B.* Back or front view: normally erect, right lateral flexion, left lateral flexion.
- V. *Position of Trunk*
  - *A.* Profile view: normally erect, ventrally flexed, dorsally flexed.
  - *B.* Front or back view: normally erect, right lateral flexion, left lateral flexion.
- VI. *Total Characteristics*
  - *A.* Grace: normally graceful, clumsy, tendency to stumble.
  - *B.* Rhythm: regular, irregular.
  - *C.* Special defects.

ERRORS INVOLVED IN THE MEASUREMENTS

In order to properly evaluate results that may be obtained by the use of the various measuring procedures described above, it is necessary to know the magnitude of the errors to which these techniques of measurements are subject. The possible types of errors may be considered under the headings of accidental or chance errors and constant errors.

1. *Accidental Errors.* Accidental errors in any procedure of measurement lead to variability in the results of a series of repeated measurements upon the same object. The magnitude of the variations likely to occur in such repeated measurements may be indicated by the calculation of the "probable error of a single measure." A small probable error signifies a high degree of precision of the particular technique of measurement used. The techniques for which we have obtained estimates of the "probable error of a single measure" are: (1) the measurements of the angle at the hip-, knee-, and ankle-joints; (2) the determinations of the angle of the foot axis, and (3) the determination of the width of the base of support.

a. Angular position at hip-, knee-, and ankle-joints. As suggested above, the determination of the "probable error of a single measure" requires a large number of repeated measurements upon the same physical object. Consequently, determination of the precision of measurement of angular position at any one of the joints would require a series of measurements upon a motion-picture frame showing the joint at one particular position. The obtained index of the precision of measurement at that particular position might not, however, adequately represent the precision obtainable with the joint at other angular positions. Determination of the precision at all positions of a joint during a cycle would be unduly arduous, and not justifiable in an exploratory series of studies. We have, therefore, attempted to obtain an estimate of the average precision of measurement at each of the joints throughout an entire cycle by means of the following procedure.

Measurement of the successive frames necessary for construction of a set of graphs was repeated three times. We thus had three measures of each position of a particular joint throughout the cycle. Each set of three measures was averaged and the deviations of each measure in the set from the average of the three was obtained. All these deviations (19 × 3 = 57) were then collected together and regarded as representing deviations of measurements made upon the same position of the joint. The probable error of a single measure was then calculated in the customary way. It is apparent that the intention of this method is to separate out the variations in the measurements due to accidental errors from the variations due to real changes in the angle at the joint. The results obtained from calculating the probable errors of measurement in this way are shown in Table 1.

TABLE 1

ESTIMATES OF AVERAGE PRECISION OF MEASUREMENT OF JOINT-ANGLE THROUGHOUT A COMPLETE CYCLE OF GAIT

Joint	Estimated probable error of measurement Right side	Left side
Hip	0.8°	0.9°
Knee	1.0°	1.1°
Ankle	0.8°	0.8°

As a check on these estimates of the average precision of measurement of joint-angle, a conventional determination was made of the precision of measurement of hip- and knee-angles at Stages I and III of the gait of a particular subject. The appropriate motion-picture frames were projected and 10 measures made of each angle being examined. The measurements for each angle were averaged and the "probable errors of single measures" were determined. The results are shown in Table 2.

TABLE 2

PRECISION OF MEASUREMENT OF JOINT-ANGLE AT SPECIFIC POSITIONS OF JOINTS FOR A PARTICULAR SAMPLE OF GAIT

Joint	Stage of gait	Angle arith. mean	Probable error of measurement
Right hip	I	193.6°	0.4°
	III	147.9°	0.5°
Right knee	I	171.7°	0.5°
	III	167.2°	0.6°

As the table shows, the probable errors determined for specific positions are of the same order of magnitude as those determined by means of the special method described above. We may, accordingly, accept the results given in Table 1 as providing a fair estimate of the precision of measurement possible in the determination of joint-angle by means of the technique described. The data on which are based the calculations of errors reported in this and the following sections were all obtained by the same observer ($G1$).

b. Precision of measurement of the angle of the foot axis. As indicated in an earlier section, the "angle of the foot axis" is determined indirectly on the basis of direct measures made of extents a, g, b, h, and f represented in Figure 6B and in the following equation:

$$\beta = \text{arc sin}\ \frac{25}{f}\left(\frac{a}{g} - \frac{b}{h}\right). \tag{6}$$

We may calculate a measure of precision for the angle if the precision measures of the five factors entering into the equation are first determined; if, for example, we first determine the probable error of a single measure for each of the factors or components.

The method,[4] then for determining "the probable error of a single measure" of the angle β consists in first determining the effect upon this angle of a given deviation in each of the components, and then determining the most probable value of all the combined effects. The effect on the angle β of a given deviation in any one of the components can be calculated by obtaining for equation (6) above the *partial derivative* of β with respect to the given component and then multiplying this result by the precision measure of that component When this procedure has been followed in turn for each of the factors in the equation, the most probable value of the combined effect of all the deviations in the factors is taken to be the square root of the sum of the squares of the separate deviations in β, due to the various deviations in the components

---

[4] For further details concerning procedures for calculation of precision measures of the dependent variable in an equation, see Goodwin (8) and Drisko (4).

These steps can be concisely summarized as follows

Given any relation $X = F\ (a, b, c, \cdot \quad \cdot\ \cdot\ )$

Let δ_a, δ_b, δ_c be the precision measures (such as the probable error of a measure) of a, b, c, etc , let Δ_a, Δ_b, Δ_c be the corresponding deviations in X produced by these deviations in a, b, c, etc , and let Δ be the resultant effect of the separate effects Δ_a, Δ_b, Δ_c, etc Then for any component k,

$$\Delta_k = \frac{\delta X}{\delta_k} \quad \delta_k$$

And the most probable value of Δ is,

$$\Delta = \sqrt{(\Delta_a)^2 + (\Delta_b)^2 + (\Delta_c)^2 + \cdot \quad \cdot\ \cdot}$$

To obtain the necessary data for using the procedures just described, the usual motion-picture records were made with one of the subjects standing in front of the camera with his right foot placed to provide an "angle of the foot axis" of approximately 10° One of the frames of these pictures was projected and 10 measurements made of the extents a, g, b, and h represented in equation (6) The measures of f were made directly on the subject's foot From these data, arithmetical means and the probable errors of the various sets of measurements were computed These values together with equation (6) provided the necessary material for applying the principles summarized above. The results of the measurements of the various factors are given in Table 3 The calculated value of the angle β was 11.8° with a "probable error of a single measure" of 88° The precision of this type of measurement thus compares favorably with the precision attained in the more direct measurement of joint-angle as reported in an earlier section

TABLE 3

DATA USED IN CALCULATIONS OF PRECISION OF MEASUREMENT OF ANGLE OF FOOT AXIS, AND WIDTH OF BASE

Factor	Determination of angle of foot axis Arith mean	 P E of a measure	Determination of width of base Arith mean	 P E of a measure
a	3 2	055	3 0	045
b	2 4	038	3 5	045
g	4 5	022	4 8	000
h	4 8	000	5 3	045
f	25 6	051	—	—

c Precision of measurement of the width of the base of support The equation expressing the dependence of the width of the base of support on directly measurable factors (see Figure 7) is

$$w = 25 \frac{b}{h} - \frac{a}{g}$$

In the case of appropriate motion-picture frames of the gait of one subject, 10 measures were made of each of the factors on the right-hand side of the above equation The results are given in Table 3

Computations carried out with these data, in the manner described previously, result in a value for the width of the base of support of 9 cm with a "probable error of a single measure" of .5 cm

2 *Constant Errors* In this section we may consider errors that may appear in measurements of the angular position at a joint due to constant factors present during certain of the measurements. Factors that may be considered here are (1) differences in the position of the subject in the field of view of the camera, such as occur, for example, in the profile pictures when the subject crosses from a position directly before the camera to a position at the extreme left or right, and (2) deviations of the plane of the leg, during the recovery stage, from a plane parallel to the direction of progression of the subject, due to abduction at the hip The position of the subject in the field of view may be specified in terms of the angle which a line connecting the subject to the camera makes with the optical axis of the camera. This angle may be referred to as the angle subtended at the camera The plane of the leg of the subject may be specified in terms of the angle lying between the plane of the leg and the plane through the leg parallel to the di-

rection of progression. By way of empirical test of the effect upon the measurement of joint-angle of the two factors mentioned, profile views of one subject were taken with the joints of the leg at certain fixed positions, and the plane of the leg, and the position of the subject in the field varied in different trials. The angles at the various joints were then measured in the usual way. The angles at which the joints were fixed for these trials were 120° for the hip, 65° for the knee, and 90° for the ankle.[5]

The results of the measurements are summarized in Table 4. The bottom row of the table shows the differences in measured joint-angle found associated with differences in the position of the subject in the field of view. Section *C* of the table shows the differences in measured angle found associated with differences in the plane of the leg. The largest difference found in any case was one of 3°. As reported earlier, the practice was adopted of regarding any angular difference less than 4° as lying within the limits of possible experimental errors, in accordance with the finding that the probable error of a single measure was about 1°. It is, therefore, unnecessary to make any special correction in the measures of joint-angle for errors due to position of the subject in the field of view, within an angle of 15°, or for angular deviations of the plane of the leg less than 22.5°. In the gaits of a series of normal subjects, deviations of the plane of the leg, in the recovery interval, ex-

TABLE 4

INFLUENCE UPON MEASUREMENT OF JOINT-ANGLE OF POSITION OF SUBJECT IN FIELD OF VIEW, AND OF PLANE OF LEG

Position of subject in field	*A* Plane of leg 0° Measured angle at			*B* Plane of leg 22.5° Measured angle at			*C* Differences 0° plane—22.5° plane		
	Hip	Knee	Ankle	Hip	Knee	Ankle	Hip	Knee	Ankle
Center (C)	125.5°	66.0°	90.0°	120.5°	65.5°	90.5°	0.0°	0.5°	0.5°
15° to left (L)	125.5°	63.0°	90.0°	121.0°	65.0°	92.0°	—0.5°	—2.0°	—2.0°
Difference (C-L)	0.0°	3.0°	0.0°	—0.5°	0.5°	—1.5°			

---

[5] For establishing these positions, we used a plumb-line goniometer which was designed for making measurements of joint-angle directly upon an individual. This instrument is described in a forthcoming paper (6).

ceeding this amount were not observed. In case errors due to constant factors, such as those discussed above, should prove under some conditions or in particular types of subjects to be of appreciable magnitude, calibration curves might be constructed which would permit a correction to be made in a given measurement to allow for the presence of the constant errors.

D Some Examples of the Use of the Present Method

The method of analysis of gait described in the present paper has been found of value in the investigation of gait in various special types of individual We have used it, for example, in tracing the changes in gait which occur in the course of normal motor development and in determining the defects in gait associated with certain

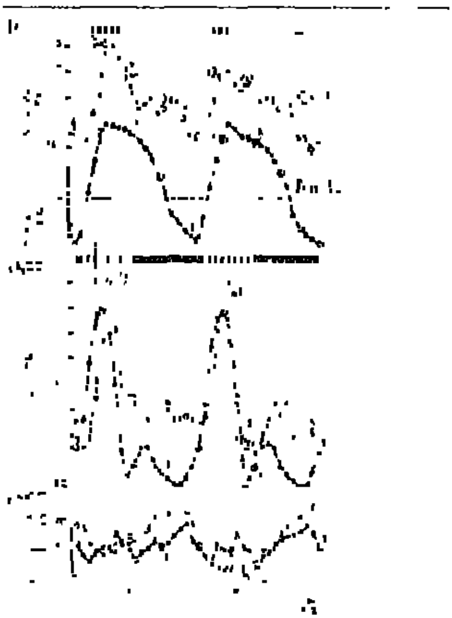


FIGURE 8

Graph of Gait of Child (assisted)

Age, 12 yrs, sex female, compared with graph for normal adult

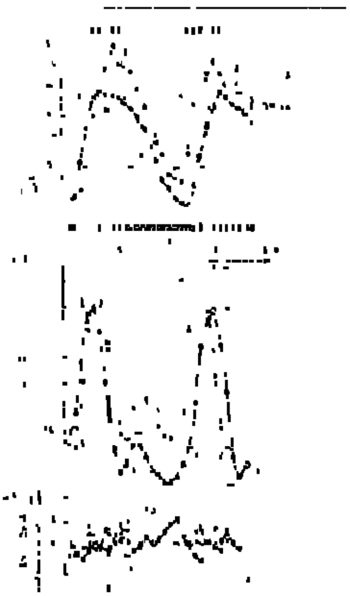


FIGURE 9
GRAPHS OF GAIT
Mongolian mental defective compared with normal adult

developmental deficiencies. The detailed results of these studies will be reported in later papers. By way of illustration, however, Figures 8, 9, and 10 are included here to show the nature of the curves obtained from the use of the method with a number of special types of individuals Figure 8 shows curves representing the gait of a child shortly before she learned to walk without support, Figure 9, curves representing the gait of an adult male Mongolian mental defective, and Figure 10, curves representing the gait of an adult subject with infantile cerebral palsy of the spastic type In each case the curve of a mentally normal adult without motor defects is included for comparison Even casual inspection of the curves reveals marked differences in their characteristics relative to those of a normal individual and indicates the sensitivity of the method in providing concrete measures of deviations from normal gait which

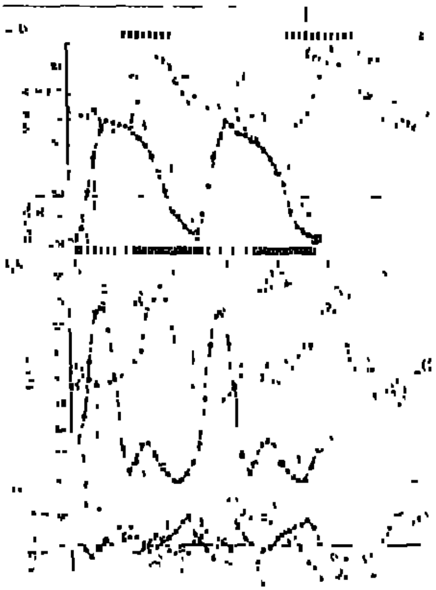


FIGURE 10

GRAPHS OF GAIT

Spastic subject compared with normal adult

may be but vaguely sensed as a result of ordinary methods of observation. More complete data on the gaits of these different types of individual will appear in later reports.

REFERENCES

1. BRAUNE, W., & FISCHER, O. Der Gang des Menschen. *Abh. d. math.-phys. Klasse d. konig. sachs. Gesellsch. d. Wiss.*, 1895, **21**, 151-322.
2. BURNSIDE, L. H. Coordination in the locomotion of infants. *Genet. Psychol. Monog.*, 1927, **2**, 279-372.
3. CARLET, G. Etude sur la locomotion humaine. *Rev. Cour. Sci.*, 1872, **3**, 476-477.
4. DRISKO, W. J. Precision of measurements. Pp. 300-303 in Standard handbook of electrical engineers, 6th ed. New York: McGraw-Hill, 1933.
5. DOUGAN, S. The angle of gait. *Amer. J. Phys. Anthropol.*, 1924, **7**, 275-279.

6 Glanville, A D, & Kreezer, G Norms for the maximum amplitude and velocity of joint movement in normal adults *Human Biol* (In press)

7 ——— Quantitative studies of human gait and its development. *Psychol Bull*, 1936, **33**, 609 (Abstract)

8 Goodwin, H M. Elements of the precision of measurements and graphical methods. New York· McGraw-Hill, 1920 Pp. 116 (See especially pp 27-30)

9 Kraus, A M The difference between a muscular and a neuromuscular interpretation of walking *Arch Neur & Psychiat*, 1923, **9**, 184-207.

10. Kreezer, G Motor studies of the mentally deficient quantitative methods at various levels of integration *Tr School Bull*, 1935, **32**, 125-135

11 McGraw, M B Growth· A study of Johnny and Jimmy New York Appleton-Century, 1935 Pp xxii + 319 (See especially pp 73-86)

12 Marey, E J Le mouvement Paris Masson, 1894 Pp vi+335

13 Mommsen, F Über den zeitlichen Ablauf der Schrittfolge und der Fussabwicklung beim Gehen und ihre Bedeutung fur die Analyse und Begutachtung von Gangstorungen *Zsch f orthop Chir*, 1932, **57**, 1-18.

14 Schwartz, R P, & Vaeth, W Method for making graphic records of normal and pathological gait *J Amer Med. Asso*, 1928, **90**, 86-89

15 Schwartz, R. P, & Heath, A L Pneumographic method of recording gait *J Bone & Joint Surg*, 1932, **14**, 783-794

16 Schwartz, R P, Heath, A L, & Wright, J N Electrobasographic method of recording gait, *Arch Surg*, 1933, **27**, 925-934

17 Schwartz, R P, Heath, A. L, Misiek, W., & Wright, J N Kinetics of human gait *J Bone & Joint Surg*, 1934, **32**, 343-350

18 Schwartz, R P, Heath, A L, & Misiek, W. The influence of the shoe on gait, as recorded by the electrobasograph and slow-motion picture. *J. Bone & Joint Surg.*, 1935, **33**, 406-418

19 Shirley, M. M The first two years Vol I Minneapolis· Univ. Minnesota Press, 1931 Pp xv + 227 (See esp. pp. 59-108)

20. Vierordt, H. Das Gehen des Menschen in gesunden und kranken Zustanden nach selbstregistrierenden Methoden dargestelt Tubingen Laupp, 1881 Pp vii + 206

21 Weber, W E, & Weber, E F. Mechanik der Menschlichen Gehwerkzeuge. Gottingen Dietrich, 1836 Pp xxvi+426

Training School
Vineland, New Jersey

THE POSTOPERATIVE EFFECTS OF REMOVAL OF THE STRIATE CORTEX UPON CERTAIN UNLEARNED VISUALLY CONTROLLED REACTIONS IN THE CAT*†

University of Rochester

KARL U. SMITH

A. INTRODUCTION

The interpretation of the maintenance of certain unlearned[1] visually determined reactions in mammals after destruction of the occipital cortex in relation to the visual sensitivity of these animals has presented a problem since the early observations of Goltz (9), Munk (25), and others of their time These experimenters observed that, after decortication or after removal of the occipital lobes in the dog, closure of the eyelids and pupillary constriction to marked changes in retinal illumination could be elicited in a manner not unlike the manifestations of these reactions in the normal animal Munk believed that the ability of the animal to blink to a bright light was no indication of vision, but Goltz held that an organism possessing this capacity could not properly be called blind.

More recently, many observers have been concerned with the neural control of the pupillary, palpebral, and other unlearned visually controlled reactions, especially in relation to the functions of the cortex, thalamus, and mid-brain

Observations have shown that cortical extirpations of any sort have no influence, either quantitatively or qualitatively, upon pupillary constriction (Byrne, 1933). However, many experimenters, among whom may be mentioned Karplus and Kreidl (10) and Parsons (26), have noted that pupillary dilation may be induced by electrical stimulation of various areas of the cerebral cortex. Kar-

---

*Received in the Editorial Office on August 5, 1936

†The experiments reported here were conducted in the Psychological Laboratory of Brown University.

[1] The term unlearned is here employed to designate the fact that the responses under consideration were present in the normal animal at the start of the experiments.

plus and Kreidl showed that, although electrical stimulation of the frontal cortex induced pupillary dilation, elimination of this part of the brain in cats did not abolish the response upon electrical stimulation of the hypothalamus. Byrne (4), in extensive experiments with cats, has established that decortication or removal of the occipital lobes does not modify pupillary dilation when it is induced by retinal illumination or by means of electrical stimulation of other parts of the body

Taking the evidence just cited and additional observations of the effects of the injection of ergotoxine upon pupillary dilation resulting from cortical stimulation, Byrne (4) argues that the cortex does not contain centers which directly mediate pupillary dilation or constriction. He holds that electrical stimulation of the frontal, occipital, and temporal areas of the cortex gives rise to pupillary dilation or constriction only indirectly, and then only under such conditions as may involve cortical stimulation of a degree sufficient to give rise to a general reverberation of nervous effect comparable to a "startle" reaction

The dependence of pupillary dilation and constriction upon connections in the thalamus, hypothalamus, and mid-brain, as indicated by experiments involving electrical stimulation, extirpation, and drug injections, has been considered in great detail in the work of Bechterew (3), Byrne (4), Ferrier and Turner (7), Karplus and Kreidl (11), Keller and Stewart (12), Magoun (20), Ranson and Magoun (27, 28), Sachs (29), and Spiegel and Nagaska (33), to which the reader is referred for extended consideration of this general topic.

The neural relations of palpebral closure and widening, as well as corresponding lens reactions, have not received the wide attention heretofore directed toward the understanding of pupillary responses Marquis (21) comes to the conclusion that, in man, reflex closure of the lids to increase in retinal illumination is permanently lost after bilateral destruction of both occipital lobes He bases his statements upon the review of nine cases of individuals blind in consequence of occipital lesions Hilgard and Wendt (8) conclude, similarly, that lid closure is mediated cortically in man, although these investigators found that this response could be elicited by shining a very bright light into the supposedly blind portion of the visual field of a patient with the right occipital lobe destroyed They explain the outcome of their observations as a result of the spread of the light to unaffected regions of the visual field

As a result of the experiments with decorticate animals (Byrne, 1933, Goltz, 1892; Munk, 1894; ten Cate, 1934), however, it seems safe to conclude that the cortex in the cat and dog does not contain fibers which are necessarily involved in the maintenance of reflex closure of the eyelids to changes in light intensity Byrne (4), whose ingenious experiments have covered the neural control of all phases of the pupillary, palpebral, and lens reactions, has emphasized their cooperation with one another in connection with the general statement that reflex adjustment of the lids together with corresponding pupillary and lens reactions are mediated through subcortical retino-neural pathways He bases this conclusion upon a quantitative analysis of pupillary, palpebral, and lens reactions in relation to decortication, various types of brain-stem transections, and the effects of drug injections.

Although much is yet to be known concerning the relation of the cortex to the response of the pupils, lids, and lenses, there is yet a greater dearth of knowledge regarding the involvement of this part of the brain in coordinated and conjugate movements of the eyes Spiegel (31) and Spiegel and Teschler (34) have shown that the lateral movements of the eyes, induced in cats by means of electrical stimulation of the frontal and occipital cortex, may be modified by bilateral injuries to the vestibular nuclei. The results of the experiments are said to indicate that horizontal movements of the eyes are abolished as a result of such injuries, whereas vertical movements are not so modified On the basis of the results of additional experiments, Spiegel and Aronsohn (32) conclude that the vestibular nuclei are intercalated in the pathway of corticofugal impulses to the efferent mechanisms of horizontal eye-movements It may be said, therefore, that the cortex does not control such movements directly. According to Marquis (21), fixation movements of the eyes are abolished in dogs after removal of the occipital lobes

Yet another aspect of the understanding of the reactions under consideration is that of their learned stimulus control and their relation to responses of the organism as a whole. Marquis (22) has been able to show in this connection that palpebral closure may be conditioned in dogs and monkeys after destruction of the occipital lobes, when the conditioned stimulus consists of a light which will not otherwise elicit the reaction In addition, Byrne (4) has stated as a result of his extensive experiments

> "The mechanisms for closure of the palpebrae, pupil and lens (accommodation) act synchronously in a co-ordinated manner with one another as well as with the basic mechanisms which mediate whatever movements or adjustments in approach—for purposes of aggression or investigation—the animal may be making, voluntarily or automatically " (p 102)
>
> "The mechanism which mediates reflex palpebral widening has, through the medium of the affective (includes the nociceptive) system, such close functional relations with the mechanisms which mediate reflex widening of the pupils and lens (accommodation for distance) that the three sets of mechanisms function as a closely co-ordinated unit" (p 110).

Excepting the statements of Byrne (4) and Marquis (22), just cited, and those of Goltz (9) mentioned at the beginning of this paper, none of the observers quoted and none in the author's recollection has laid emphasis upon the fact that unlearned ocular responses, either as they may be elicited in the normal animal or in the animal without visual or neocortex, indicate certain psychological capacities of the animal to respond to external stimulation The investigation of these reactions in animals lacking cortical connections of the visual system should, therefore, lead to a better understanding of the functions of the striate cortex and of the organization of the central optic system in respect to different psychological adjustments

B EXPERIMENTS

1 *Observations* Three different series of experiments were made involving the observation in normal animals and in animals lacking the visual cortex of (1) palpebral closure to changes in light intensity brought about by movement of the experimenter's hand before the animals' eyes, (2) compensatory movements of the eyes

in a rotating visual field,[2] (3) visual placing reactions of the forelimbs under different stimulus conditions, including circumstances of stimulation which may be said to be comparable to those utilized in the second series of experiments Observations were made at periods of one week, two weeks, three weeks, three months, and five months after the operations, using different animals at the different periods of examination

2 *Operations.* The extent and character of the visual cortex of the cat has been described in detail by Alouf (1), Campbell (5), Koppen and Lowenstein (13), Winkler and Potter (36), and others, among whom there is a close agreement as to the anatomical limits of this area The precise delimitation of the area by observable landmarks on the surface of the brain permits, at the time of the operations, a determination of the extent of the visual cortex removed

Bilateral operations were performed in one stage on five animals (Subjects 17, 29, 36, 37) and in two stages on one animal (Subject 30), under completely aseptic conditions In an additional control kitten (Subject 43), the striate cortex was removed on one side only The operations were carried out under Nembutal anesthesia (Pentobarbital sodium[3]) After adequate exposure of the occipital lobes by approaches through the skull under the temporal muscles, an excision was made along the lateral sulcus, well within the boundaries of the second external convolution, extending from the vicinity of the Sulcus ansatus beyond the posterior termination of the lateral sulcus. The cut with the cautery was directed in such a manner as to sever the optic radiations. Then, using a small blunt spatula and bent-nose forceps, the excised lobe was carefully dissected out, care being taken to observe, as the lobe was removed, whether or not the dissection included, on the medial side of the brain, the Sulcus splenialis along its entire length The piece of tissue was thereafter removed and fixed in formalin for further examination Although complete anatomical controls concerning the operations are not yet available, the nature of the extirpations can be understood by reference to Figure 1, in which are repro-

---

[2] Dr D G Marquis has reported similar observations before the Psychological Colloquium of Brown University on reactions in dogs

[3] Abbot Laboratories, North Chicago, Illinois

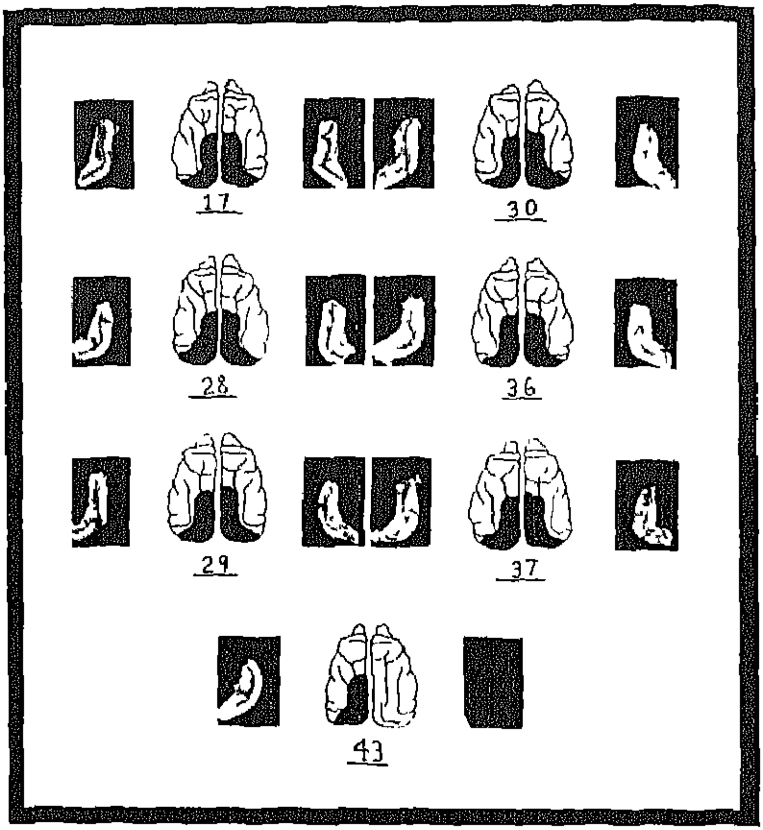


FIGURE 1

The Dissected Occipital Lobes of the Different Subjects Used in the Experiments

The lobes were removed intact in all animals except that of the right hemisphere of Subject 28, the posterior portions of which were broken off during the operation. Subject 43 was only 28 days old when the left occipital lobe was removed.

duced in actual photograph as well as in diagram the dissected occipital lobes of the different animals employed in the experiments. The diagrams, which have been superimposed upon a drawing of a normal brain, are intended only as approximations of the extent of tissue removed.

3 *Subsidiary Observations of the Behavior of the Animals with Complete Bilateral Removal of the Striate Cortex* All of the animals with bilateral operations employed in the present experiments portrayed characteristic defects of visual capacity, which have been noted by previous investigators in the dog, cat, and monkey after removal of the visual areas Subsequent to the operations the animals employed here were able to avoid objects placed in their path only by coming into direct and abrupt contact with such obstacles Under normal conditions of illumination a quick movement of the experimenter's hand directly before their eyes failed to evoke the quick retraction of the head and the reflex closure of the eyelids so typical of the visual behavior of the normal animal under such circumstances They were found to be unable to locate food quickly and efficiently when it was placed in their near vicinity, although effective behavior of this sort could be learned if the food was located day after day in the same spot within their individual quarters None of these animals could be induced to climb stairs by means of any sort of incentive, nor would they descend from an elevated surface if the conditions did not permit them to touch the floor or to gain foothold along the sides of the platform upon which they were placed Thus, animals placed on a low table, 25 cm in height, with the legs set back from the edge so as not to afford any sort of foothold, would repeatedly try to touch the floor in order to descend, but, failing, would remain in that predicament until removed

Normal animals, and the single kitten with the striate cortex removed on one side only, failed to show any defects of the sort just described, nor have such disturbances in behavior ever been observed for more than a few days in animals with extensive but incomplete removal of the visual areas These observations agree in all details with the picture of cortical blindness resulting from complete removal of the striate areas in cats and dogs described by Marquis (22), Minkowski (23, 24), ten Cate (35), and others Hence, the responses hereafter described are understood as being correlated with typical symptomatic defects consequent upon complete removal of the visual cortex, which have been observed by prior investigators.

4 *Apparatus and Procedure* In making observations of the responses of the eyelids to alteration in light intensity, the experi-

menter sat in a chair directly in front of a strong light source Intense illumination was obtained by placing a 500-watt bulb in a large parabolic aluminum reflector that was directed toward the animal as it was held in the experimenter's lap. In each case a distance of 72 cm. intervened between the surface of the cornea of the cat's eye and the light source. After quieting the animal by stroking its head, the light was turned on and the action of the pupils and lids observed. With continued adaptation to this illumination, the lids were found to be moderately closed and the pupils reduced to a width of 2 to 3 mm After adaptation to the continued illumination had taken place, an assistant quickly passed his hand between the light source and the cat's eyes at a distance of approximately 10 to 15 cm in front of the head of the animal As a control observation, a piece of cardboard, 10 cm wide and 25 cm long, was also moved in a vertical plane between the animal's head and the light source at the same distance in front of the eyes

Forced lateral movements of the eyes controlled by external stimulation may be elicited in cats by rotating a striated black and white pattern around the animal, according to the method employed by Schlieper (30) in the investigation of color sensitivity of invertebrates (Figure 2) The pattern employed in these experiments consisted of black paper strips, 5 12 cm wide, placed vertically at equal intervals of 5.12 cm upon white oilcloth, which was stretched uniformly upon a large cylindrical frame, 110 cm high and 122 cm. in diameter. The bottom of the drum as well as a rim, 20 cm. wide, around its top were also covered with the black and white stripes

A heavy steel shaft, attached to the under side of the drum at its center, rotated in a bearing mounted upon a heavy iron base. The drum could thus be rotated uniformly by means of a motor and pulleys fixed to the shaft of the motor and to that of the rotating drum. Suspended to a beam above the drum and extending downward 100 cm within it was an iron rod, 2.56 cm in diameter, on which was clamped a small wooden platform, which was so located as to permit the adjustment of the animal's head at a point equidistant from the sides of the cylinder. The animal was restrained in a box-holder and so placed upon the platform that its head was held in a horizontal position. During the experiments, the cat, re-

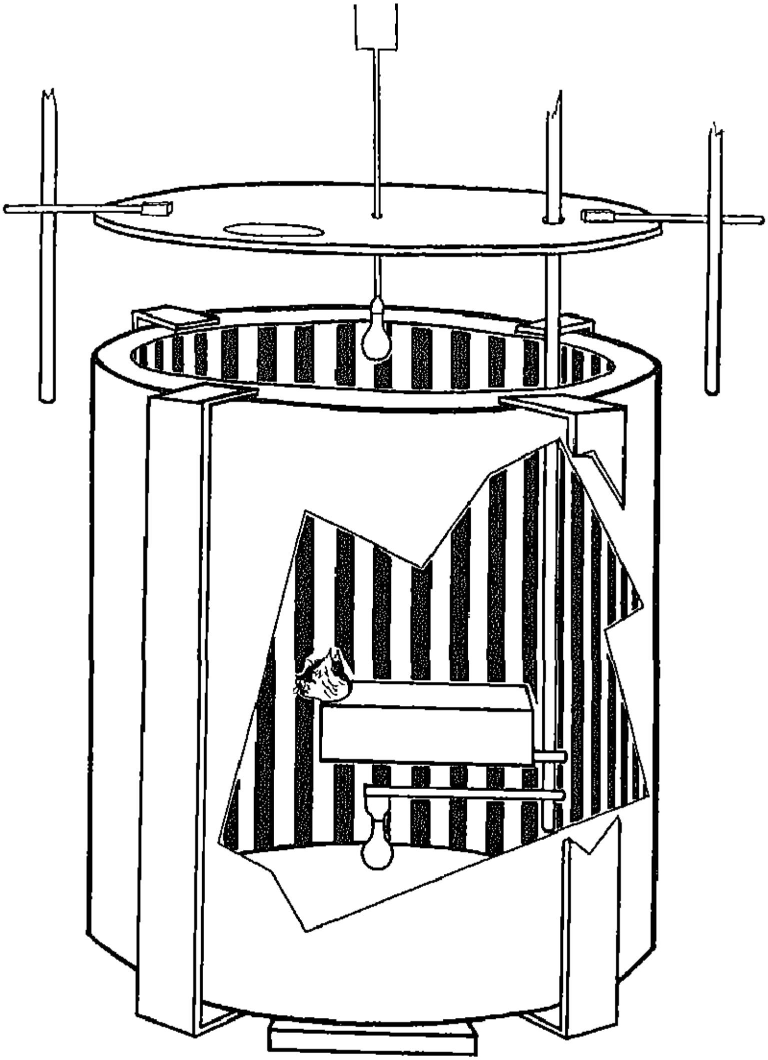


FIGURE 2
THE APPARATUS USED IN THE OBSERVATION OF COMPENSATORY EYE-MOVEMENTS TO A ROTATING STRIATED PATTERN

strained in the box-holder, was lowered to a point 90 cm below the level of the top of the cylinder, and the pattern on the sides of the drum was thereafter rotated entirely independently of the animal itself

The opening in the top of the drum was covered, during experimentation, by a rigid, white, circular screen, which contained an aperture for the rod suspending the small platform It also contained an aperture for making observations of the movements of the animal, either by means of motion pictures or by means of direct observation This screen could be raised and lowered in order to permit the placing of the animal upon the platform as well as removing the cat after each experimental setting The inside of the apparatus was illuminated by means of two 200-watt bulbs, one mounted at the center of the screen above the rotating drum and the other placed at the center of the bottom of the drum Observations on three normal animals, one kitten with unilateral removal of the striate cortex, and five animals with complete removal of the visual areas were supplemented in every case by motion pictures taken through the aperture in the screen above the drum

During the examination of visual placing reactions of the forelimbs, the experimenter grasped the cat beneath the thorax and around the hind quarters and moved it slowly toward the top surface of a table Two types of visual placing reactions were tested. (1) those ordinarily elicited by moving the animal in a horizontal direction toward a surface, and (2) those ordinarily elicited by lowering the animal in a vertical direction, head downward, toward a surface.

Experiments were carried out under three different circumstances of visual stimulation In a first series of tests the reactions were observed under ordinary circumstances of stimulation, such as that obtained at midday without the aid of accessory lighting. In these tests the animals were examined by moving them toward the top surface of a table, measuring 108 by 72 cm., which presented a completely white surface. In a second series of tests, attempts were made to elicit both the horizontal and vertical placing reactions when the top of the table was covered with black cardboard stripes 5.12 cm. in width. The stripes were firmly attached to the top of the table in such a manner that equal distances of 5 12 cm. separated the stripes In a third series of observations the top of the table exhibited a black- and white-striped pattern of lines 10 24 cm in width In the second and third series of observations, the table was illuminated by two 500-watt lights placed in large, parabolic aluminum reflectors, which were located 100 cm above the opposite ends of the table Observations of two normal animals, one kitten with complete unilateral removal of the striate cortex, and six animals with complete removal of the visual areas were supplemented in each case by motion-picture records taken during the course of the examinations.

5 *Results*

a Palpebral responses induced by objects moving across the visual field. Two normal animals and one animal with unilateral removal of the striate cortex exhibited, under conditions of daylight illumination, reflex closure of the eyelids when the experimenter quickly moved his hand or other objects across their field of vision They also gave such responses under experimental conditions in which their eyes were adapted to high illumination The reaction was elicited when the movement of objects, either the hand of the experimenter or a piece of cardboard, took place 10 to 25 cm in front of the corneal surfaces Five animals with complete bilateral removal of the striate cortex failed to respond in a similar way under daylight illumination at the different intervals of time after the operations during which the reactions were observed.

Under the conditions of intense illumination in which the animal was held at a distance of 72 cm in front of a 500-watt source, shining directly upon its head, movement of the experimenter's hand across the visual field evoked in the operated animals reflex closure of the eyelids in response to the change in retinal illumination brought about by the movement The reaction could be elicited only when the hand was passed at distances of 10 to 15 cm in front of the animal's eyes, and, when so elicited, consisted in complete closure of the eyelids which were already partly narrowed in consequence of the heightened illumination The response is the same as that observed when a strong light source is quickly presented to the cat after dark-adaptation or after the animal has become accustomed to normal conditions of illumination, and, similarly, disappears with repeated stimulations

Two control observations point to the fact that the response was determined by changes in retinal illumination rather than by changes in tactual stimulation induced by air currents set up by movement of the hand in front of the animal's head. On the one hand, the failure of the response to appear in the operated animals without visual cortex under normal conditions of illumination shows that such secondary sources of stimulation were not involved in the elicitation of the response On the other hand, the movement of the card across the visual field under the conditions of heightened illumination, an operation which reduced to a minimum these secondary sources of stimulation, evoked the response under the circumstances described.

b Compensatory movements of the eyes in a rotating visual field In the normal adult cat, rotation of the visual field, according to the method described in the foregoing section, gives rise to forced lateral movements of the eyes in the direction of the rotation, followed by a quick recovery movement in the direction opposite that of the rotation. The question as to whether or not such adjustments can be considered as true nystagmic movements is not a direct concern of the present discussion, although the fact may be pointed out that, after rotation of the field has ceased, no aftereffects of rotation could be ascertained In addition, the responses were observed to cease after a few minutes of rotation, the cat thereafter falling into a lethargic sleep-like state with eyes partly or completely closed.

Using a speed of rotation of approximately 3 r p m , a single animal with unilateral removal and five animals with complete bilateral removal of the visual cortex were observed to exhibit forced lateral movements of the eyes which could not be differentiated from those of normal animals. Reversals of rotation gave rise to reversals in the direction of the movements, and increase in the speed of rotation up to a certain point was found to produce noticeable increase in the speed of the movements Reduction in the level of the illumination of the visual field by utilizing two 50-watt instead of two 200-watt lights in the apparatus failed to eliminate the response in two animals with bilateral removal of the striate cortex No observations were made with lines of a width greater or smaller than 5 12 cm , although the method here employed offers a means of such variation in the measurement of the visual acuity of animals so prepared The observations were supplemented by motion-picture records of the responses found.

c Visual placing reactions of the forelimbs The observations made in this study upon the effect of removal of the visual cortex upon visual placing reactions constitute an extension of the experiments of Bard and Orias (2), in which it was found that such operations eliminate these reactions in the cat when tested under normal conditions of illumination In the present experiments, five animals with complete bilateral removal of the visual cortex failed to exhibit either horizontal or vertical placing reactions under conditions of daylight illumination, or when moved toward the top surface of a table which was highly illuminated Nor were these reactions evoked when the animals were moved towards a striated surface which was similar to that utilized in the experiments with the rotating visual field. In comparison, one animal with unilateral removal of the striate cortex exhibited these reactions under all conditions of illumination and under all circumstances in which different surface patterns were employed on the table

C DISCUSSION

On the basis of an objective interpretation of the responses observed in the present experiments, the following conclusions may be drawn concerning the visual capacity of the cat following removal of the visual cortex

The first series of observations on palpebral reactions demonstrate that, after complete extirpation of the striate cortex in cats, there are maintained differential responses of the eyelids which may be proved to be functionally related to changes in visual intensity induced by movement of objects across the visual field The observation that these animals give avoidance reactions to movement of a card or of the experimenter's hand across a highly illuminated field is significant in showing that changes in retinal illumination, so induced, may serve as a basis for avoidance reactions to "objects" or to "threatening gestures" Accordingly, it is concluded that cats lacking the visual cortex possess a rudimentary capacity to avoid "objects" or "threatening gestures," an interpretation which is in no way contradicted by the fact that the responses upon which this capacity is based are predominantly controlled by changes in light intensity. That no such ambiguity of interpretation exists is shown by the fact that, in normal animals, generalized reactions in avoidance of objects may derive directly from changes in retinal illumination as well as from differences in visual size, color, pattern, or movement The difference between the results obtained here and those secured by previous experimenters who have failed to observe similar responses in the dog after removal of the visual areas is explained by the heightened conditions of illumination utilized in the present experiments

Failure of destruction of the striate areas to abolish compensatory eye-movements to a moving striated pattern points to the fact that these reactions are subcortically controlled Their functional correlation in the operated animals with the speed and direction of the rotating pattern permits the conclusion that the nervous mechanism by means of which these responses are determined, are also capable of mediating impulses associated with movement and pattern vision. Since the localizing reactions of the eyes in the normal animal are considered as evidence of "object" vision when such reactions are based upon fixed or moving patterned contours in the field of vision, the responses observed here in cats without striate cortex may also be taken as evidence for a type of object perception

The experiments made here confirm the observations of Bard and Orias (2), who found that placing reactions of the forelimbs in the cat are abolished by removal of the striate areas According to the present investigation these responses are apparently lost not only under normal conditions of illumination but also in relation to various kinds of patterned stimulation under circumstances of increased general illumination Striated patterns, which, when rotated around the head of the animal, evoke forced lateral deviations of the eyes, fail to elicit placing reactions of the forelimbs when the animal is moved horizontally or vertically in their direction. These experiments, therefore, indicate a critical division of function between the neural mechanisms mediating localizing ocular reactions and those determining the visually controlled orientations of the forelimbs and trunk

Additional significance may be attached to the data here secured by considering the relation between the results of the present experiments and those obtained in investigations of the visually controlled discriminative behavior of mammals without the striate cortex The experiments of Lashley (14, 15, 16, 18) and of Marquis (21) with the rat and dog have shown that mammals deprived of the visual cortex lose, but can recover, the capacity to discriminate intensity differences in vision Marquis failed to find any marked increase in the threshold for the discrimination of two lights in dogs after complete removal of the striate areas, but Lashley observed somewhat more significant threshold modifications in rats A question arising out of these experiments has been What is the psychological character of the residual visual function in mammals lacking *both* of the projection areas of the visual system? Lashley (17) and Lashley and Frank (19) attempted to answer this query in experiments with rats, in which it was shown that animals completely deprived of the striate areas permanently lost the ability to discriminate differences in visual patterns under circumstances of daylight vision The immediate conclusion of the experiments of Lashley and Frank has been, therefore, that mammals, after abolition of the cortical connections of the central visual pathways, entirely lose pattern discrimination but retain a crude capacity for brightness vision, although Lashley (18) himself has made certain reservations regarding such an hypothesis, because there seems to be a demonstrable relation between intensity discrimination and pattern vision It is very difficult to see, for example, how an animal can localize two lights in the field of vision without being able to distinguish contours of light, a capacity ordinarily defined as pattern vision Marquis, however, attributes the maintenance of brightness discrimination in his animals to a primitive type of light vision and believes that such defects as the inability to avoid obstacles, the failure to respond to threatening gestures under certain conditions of illumination, etc, are a direct result of a specific loss in what he calls "object vision" This capacity is assumed to be associated with the functions of the cerebral cortex as it has become elaborated in animals such as the dog in mammalian phylogenetic development

Properly considered, the present experiments bring forward certain points of evidence contrary to the view that only the cortical processes of the visual system are capable of mediating responses to changes in pattern or to the location of objects Rather, the data here secured seem to show that the subcortical neural pathways controlling the reactions of the eyelids and of the eye, as a whole, are differentiated to such a degree as to direct adjustments in conjunction with changes in pattern and movement in the visual field. These statements do not mean, of course, that pattern vision may not be lost in the operated animals in relation to discriminative orientations under certain conditions of general illumination They do mean, however, that different responses must be considered in attempting to understand the neural basis of pattern, movement, and object vision. So far as the writer is aware there is available no explanation of the apparent functional separation between the neural mechanisms of the ocular reactions and those of visual placing reactions to changes in pattern and movement Evidently, as shown here by the experiments on compensatory eye-movements and visual placing reactions, this division of function cannot be attributed to the differences in stimulation employed in testing these reactions, although subsequent experimentation may prove this to be in part the case in respect to other similar conflicting data

The foregoing discussion raises a question as to the relation between the subcortical mechanisms of the ocular reactions and the functions of the visual cortex in regard to pattern, movement, and object vision According to the results obtained in this study, the striate cortex of the cat seems to constitute a mechanism of more elaborate differentiation for these particular psychological adjustments which are mediated also in a simpler way at other levels of the visual system There is no foundation in fact for the view, however, that the functions of pattern, movement, and object vision have been entirely "corticalized" in phylogenetic development up to and including animals such as the cat In respect to the organization of the central visual system in this animal, the present evidence seems to prove that the subcortical processes function in cooperation with cortical mechanisms in the mediation of pattern, movement, and object vision The subcortical visual processes in the cat maintain, in part at least, the localizing functions (responses to pattern, movement, and objects) mediated through the mid-brain and thalamus at more primitive levels in phylogeny.

In the foregoing interpretations, the question has not been raised as to whether or not the visually controlled palpebral reactions, compensatory eye-movements, and placing reactions of the forelimbs in the normal and operated cat involve phenomena which are said to be described by the expression "conscious visual experience." As far as the writer is aware, there is at present no method of observation available by means of which an answer to this question can be secured It should be noted, in addition, that the problem of the relation between "consciousness" and behavior can be considered as distinct from, and, indeed, entirely beyond the scientific questions of the stimulus and neural control of behavior in phylogenetic and ontogenetic development (Carmichael, 1933). Similarly, according to the opinions of the present writer, it seems possible to view as insoluble the original problem raised by Goltz (9) and Munk (25) concerning whether or not the unlearned ocular reactions are a true indication of "visual experience" in mammals lacking the cortical pathways of the optic system In regard to adjustments based on visual stimulation in cats lacking the striate cortex, however, the present results seem to prove that these ocular reactions involve stimulus-response relations which may be defined in objective terms as movement, pattern, and object vision With reference to the visual placing reactions of the forelimbs in the same animals, similar relations are not demonstrable

D SUMMARY AND CONCLUSIONS

The experiments may be summarized as follows

1 Complete removal of the visual cortex in six cats was followed by the appearance of typical defects in the ability to avoid obstacles, to locate food, to climb stairs, and to descend from elevated surfaces on the basis of visual cues These deficiencies of behavior were found over the entire period of time during which the animals were under observation.

2. Although closure of the eyelids to movement of a card or of the experimenter's hand under normal conditions of illumination could not be elicited in five animals, such responses were evoked in all five when the card or hand was moved across the visual field of the animal under extremely high retinal illumination

3 Forced lateral deviations of the eyes to a rotating striped pattern were observed in five of the animals after the operations

4 Visual placing reactions of the forelimbs were abolished in six animals under circumstances of stimulation similar to those employed in testing eye-movements

5 The failure of the bilateral removal of the striate areas to abolish the avoidance reactions of the eyelids to movement of the experimenter's hand across the visual field under heightened conditions of illumination is taken as evidence for a rudimentary type of "object" perception in these animals.

6 The failure of the bilateral removal of the striate areas to abolish forced lateral movements of the eyes to a rotating striated pattern points to the existence of movement and pattern vision in cats lacking the visual cortex

7 The abolition of visual placing reactions of the forelimbs to different striped surfaces indicates critical division of function between the nervous mechanisms of pattern vision mediating ocular reactions and those controlling the visually determined responses of the trunk and forelimbs

8. In respect to pattern, movement, and object vision in the cat, the subcortical neural mechanisms of the ocular reactions seem to function in cooperation with more highly differentiated cortical processes The subcortical neural pathways of the central visual system in the cat retain in part the localizing functions which are mediated entirely through the mid-brain at more primitive levels in phylogeny.

REFERENCES

1 ALOUF, I Die vergleichende Cytoarchitektonic der Area Striata *J Psychol u Neurol*, 1929, **38**, 5-41

2. BARD, P, & ORIAS, O. Localized cortical management of visual and labyrinthine placing reactions *Amer J Physiol*, 1933, **105**, 2-3

3. BECHTEREW, W. Functionen der Nervencentra Vol II Jena Fischer, 1909. Pp viii+693-1336

4. BYRNE, J G Studies on the physiology of the eye London Lewis, 1933 Pp xii+428

5 CAMPBELL, A. W Histological studies of the localization of cerebral function Cambridge Cambridge Univ. Press, 1905 Pp xx+428

6 CARMICHAEL, L Origin and prenatal growth of behavior In *A handbook of child psychology*, 2nd ed rev, ed by C Murchison. Worcester, Mass Clark Univ Press, London Oxford Univ Press, 1933 Pp. 31-159

7 Ferrier, D, & Turner, W A Experimental lesions of the Corpora Quadrigemina in monkeys *Brain*, 1901, **24**, 27-46

8 German, W J, & Fox, J C (In collaboration with F H Couch, E R Hilgard, W R Miles, R R Sears, P K Snedden, & G R Wendt) Observations following unilateral lobectomies In *Localization of function in the cerebral cortex Proc Asso Res Nerv & Ment Dis*, 1932, **13**, 378-434

9 Goltz, F Der Hund ohne Grosshirn *Pfluger's Arch f d ges Physiol*, 1892, **51**, 570-614

10 Karplus, J P, & Kreidl, A Gehirn und Sympathicus II Ein Sympathicuszentrum im Zwischenhirn *Pfluger's Arch f d ges Physiol*, 1910, **135**, 401-416

11 ——— Über die Bahn des Pupillereflexes *Pfluger's Arch f ges Physiol*, 1913, **149**, 115-155

12 Keller, A. P, & Stewart, L The superior colliculus and the pupillary light reflex in the cat *Amer J Physiol*, 1932, **101**, 64

13 Koppen, M, & Lowenstein, S Studien uber den Zellenbau der Grosshirnrinde bei den Ungulaten und Carnivoren und uber die Bedeutung einiger Furchen *Monatschr f Psychiat u Neurol*, 1908, **18**, 481-509

14. Lashley, K S Studies of cerebral function in learning IV Vicarious function after destruction of the visual areas *Amer J Physiol*, 1922, **59**, 44-71

15 ——— Studies of cerebral function in learning VII The relation between cerebral mass, learning, and retention *J Comp Neur*, 1926, **41**, 1-58

16 ——— The mechanism of vision II The influence of cerebral lesions upon the threshold of discrimination of brightness *J Genet Psychol.*, 1930, **37**, 461-480

17. ——— The mechanism of vision IV The cerebral areas necessary for pattern vision in the rat *J Comp Neur*, 1931, **53**, 419-478

18 ——— The mechanism of vision XII Nervous structures concerned in the acquisition and retention of habits based on reactions to light *Comp Psychol Monog*, 1935, **11**, 43-79

19 Lashley, K S, & Frank, M The mechanism of vision X Postoperative disturbances of habits based on detail vision in the rat after lesions in the cerebral visual areas. *J. Comp Psychol*, 1934, **17**, 355-393

20 Magoun, H W Maintenance of the light reflex after destruction of the superior colliculus in the cat *Amer J Physiol*, 1933, **111**, 91-98

21 Marquis, D G Effects of removal of the visual cortex in mammals, with observations on the retention of light discrimination in dogs In *Localization of function in the cerebral cortex Proc Asso Res Nerv. & Ment Dis*, 1932, **13**, 558-592

22 ——— Phylogenetic interpretation of the functions of the visual cortex *Arch Neur & Psychiat*, 1935, **33**, 807-815

23. Minkowski, M. Zur Physiologie der Sehsphare. *Pfluger's Arch. f. d. ges. Physiol.*, 1911, **14**, 171-327.

24. ——— Experimentelle Untersuchungen uber die Beziehungen der Grosshirnrinde und der Netzhaut zu den primaren optischen Zentren, besonders zum Corpus Geniculatum externum. *Arb. aus dem hirnanat. Inst. in Zurich.* Wiesbaden: Bergmann, 1913, **7**, 255-362.

25. Munk, H. Ueber den Hund ohne Grosshirn. *Arch. Anat. u. Physiol., Physiol. Abt.*, 1894, 355-369.

26. Parsons, J. H. On dilation of the pupil from stimulation of the Cortex Cerebri. *J. Physiol.*, 1901, **26**, 366-379.

27. Ranson, S. W., & Magoun, H. W. Respiratory and pupillary reactions induced by electrical stimulation of the hypothalamus. *Arch. Neur. & Psychiat.*, 1933, **29**, 1179-1194.

28. ——— The central path of the pupilloconstrictor reflex in response to light. *Arch. Neur. & Psychiat.*, 1933, **30**, 1193-1204.

29. Sachs, E. On the structure and functional relations of the optic thalamus. *Brain*, 1909, **32**, 95-186.

30. Schlieper, C. Farbensinn der Tiere und optomotorische Reaktionen. *Zsch. f. vergl. Physiol.*, 1927, **6**, 453-472.

31. Spiegel, E. A. Role of the vestibular nuclei in the cortical innervation of the eye muscles. *Arch. Neur. & Psychiat.*, 1933, **29**, 1084-1097.

32. Spiegel, E. A., & Aronsohn, L. Interaction of cortical and labyrinthine impulses of ocular muscle movements. *Amer. J. Physiol.*, 1934, **109**, 693-703.

33. Spiegel, E. A., & Nagaska, G. Experimentalstudien am Nervensystem. VI. Über die Beziehung des Pupillen-reflexbogens zum vordern Vierhugel. *Pfluger's Arch. f. d. ges. Physiol.*, 1927, **215**, 120-132.

34. Spiegel, E. A., & Teschler, L. Experimentalstudien am Nervensystem. XII. Über die Beziehung der Blickbahn zu den Vestibulariskernen. *Pfluger's Arch. f. d. ges. Physiol.*, 1929, **222**, 359-370.

35. ten Cate, J. Akustische und optiche Reaktionen der Katzen nach teilweisen und totalen Extirpationen des Neopalliums. *Arch. Neérl. Physiol., l'Homme & Anim.*, 1934, **19**, 191-264.

36. Winkler, C., & Potter, A. An anatomical guide to experimental researches on the cat's brain. Amsterdam: Versluys, 1914. (30 plates with accompanying descriptions.)

University of Rochester
Rochester, New York

A FURTHER STUDY OF AFFECTIVE FACTORS IN LEARNING*

University of California

H. D. Carter and H. E. Jones[1]

Introduction

In an earlier report (2) of a learning experiment conducted with elementary and junior high school children, it was shown that reliable differences occurred in the immediate recall of affectively toned words. Words which the children classified as characteristically *pleasant, indifferent,* or *unpleasant* were chosen for the experiment. Learning was relatively easy for pleasant words and relatively hard for indifferent words. The unpleasant words were of intermediate difficulty in learning. The present report deals with results from comparable procedures applied to an older age group.

The Data

Individual learning tests requiring about a half hour for administration were given to each of 100 college students. In the group tested there were 56 males and 44 females. The age range was from 17 to 25 years, with a mean of 20.8 years and a standard deviation of 1.76 years. There were 12 freshmen, 18 sophomores, 32 juniors, 27 seniors, and 11 graduate students. As compared with the relatively unselected group of adolescent children tested in the preliminary study, these college students are, on the average, eight years older and represent a slightly higher selection in intelligence and social status.

The words used in the learning tests were as follows:

> Pleasant words (*P*): *mother, father, candy, movies, love, marry, dance,* and *kiss*.
>
> Indifferent words (*I*): *pencil, enter, notice, glass, turn, fill, middle,* and *lid*.

---

*Received in the Editorial Office on August 24, 1936.

[1] The writers are indebted to Mr. Charles E. Buffinger and Miss B. G. Boucher for assistance in the data collection, and to Dr. H. S. Conrad for criticism of the manuscript.

Unpleasant words (*U*): *pimple, fight, bleed, fear, insult, garbage, cheat,* and *vomit*

These words had originally been chosen from larger lists because they were familiar to all the adolescents in the original study, and because the words in the three categories were uniformly rated by the children as pleasant, indifferent, and unpleasant, respectively. The selection of the words and later evaluation of them as emotional stimuli are more fully discussed in the preliminary study (2). In that report it was shown that the unpleasant words produced largest deflections of the galvanometer, the pleasant words somewhat smaller deflections, and the indifferent words produced smallest deflections. Thus, both ratings and measures of physiological response are available as evidence of the potency of the words as emotional stimuli. The procedure of rating the words and of obtaining galvanometric records was not repeated in the present group.

The learning task employed was one of immediate recall. The words were learned by a paired-associates method, in which the task was to recall the words in response to pictures. The pictures were a series of conventional, easily duplicated illustrations of persons taken from a clothing catalogue. Three trials were given after the initial exposure, with prompting on each trial if the recall was incomplete or incorrect. The precautions taken to eliminate common sources of error have been described more fully elsewhere (2, 3) and will be mentioned only briefly. The words in the three categories were evenly distributed throughout the list; at the beginning and at the end of the series, three pairs of picture-word associations were used as buffers, in order to eliminate effects of the tendency to favor words at the beginning and end of the series. Careful rotational procedures were employed as a further precaution.[2] Because of the rotation, subjects associated the particular words with different pictures, hence no particular cues relating to individual word-picture combinations could affect the group-learning results. In the initial learning series, a picture was exposed for 3 sec., followed by 2 sec.

---

[2] The group data reveal almost no position effects, even for the buffer words at the beginning and end of the series; this is probably because the associated pictures offered better cues for learning. However, the controls exercised are known to be desirable, in view of individual differences which have been observed.

in which both the picture and the associated word were exposed In the recall or relearning series, a picture was exposed for 3 sec and the subject given an opportunity to recall the word, if he failed, the word was exposed again as in the initial learning

TREATMENT OF DATA AND RESULTS

Learning scores were computed for groups of pleasant (P), indifferent (I), and unpleasant (U) words separately, by summation of correct responses for items in the three categories, respectively This was done for odd and even groups of words within each category, as well as for each of the three total groups. Reliability coefficients were obtained from the correlation of odd and even halves, by application of the Spearman-Brown prophecy formula

In Table 1 are given the means, standard deviations, reliability coefficients, and intercorrelations of the learning scores for P, I, and U groups of words The intercorrelations are definitely lower than the reliabilities, but are sufficiently high to have marked effect upon the standard errors of the differences between means Corrected for attenuation, the intercorrelations would not be perfect[3] The reliability coefficients indicate that a single test is unsuitable for individual diagnosis, but the data are adequate for the study of differences in average scores It may be noted that in the previous studies of children, involving a succession of five experiments, trends

TABLE 1

MEANS, STANDARD DEVIATIONS, RELIABILITY COEFFICIENTS, AND INTERCORRELATIONS OF P, I, AND U SCORES FOR LEARNING

	P-scores	I-scores	U-scores
N	100	100	100
Mean	16 84	12 82	15 82
S D	4 13	5 19	4 11
Correlations			
P-scores	*67*	48	53
I-scores		*71*	45
U-scores			*64*

Note The reliability coefficients are given on the diagonal, in italics which are apparent in any single half-hour test are supported when more reliable data are obtained from cumulating the results of several testings. It is apparent from the means given in Table 1 that (as in the case of the data for younger adolescents) the pleasant words are learned best, the unpleasant are intermediate, and the indifferent are least efficiently learned.

---

[3] Evidence secured elsewhere, from repeated testing of the adolescent group, indicates that when scores become very reliable the intercorrelations still remain somewhat below the reliability coefficients The results apparently arise from individual differences in the extent to which emotional factors interfere with learning

Reliability of Differences. In Table 2 are given the differences between means, the standard errors of the differences, and the critical ratios. It is apparent that the indifferent words are learned significantly less well than either the pleasant or the unpleasant. The difference between the learning scores for pleasant and unpleasant words is smaller, but the critical ratio of 2.55 here indicates a high probability (99 chances in 100) that the obtained result is in the true direction. Further confidence in these findings is justified in view of their close agreement with the results from the adolescent group (2, 3).

TABLE 2

STATISTICAL RELIABILITY OF THE DIFFERENCES BETWEEN MEANS*

	Diff	$S.E._{diff}$	$\frac{Diff}{S.E._{diff}}$
P-I	4.02	484	8.31
P-U	1.02	399	2.55
U-I	3.00	496	6.04

*The *P*-scores were highest, *U* scores were second, and *I*-scores were lowest.

DISCUSSION

The results reported in this paper derive their significance largely from the fact that they constitute a confirmation, in a more mature group of subjects, of the observations made in an intensive study of learning among children. The similarity of quantitative results, both as to the order (*P, U, I*) of learning and the relative magnitude of the three differences, is striking.

In follow-up work with the adolescents (3), the experiment has been repeated with different samplings of words. The trends of results, in terms of averages, have been clearly demonstrated to be independent of the particular sampling of pleasant, indifferent, and unpleasant words.

The present results have incidental bearing upon one of the stock criticisms of work in this field. As an explanation of such learning trends, some critics have called attention to the possible importance of differences in the familiarity of the words chosen to represent the three affective categories. Thus, the pleasant words may be better remembered because they are more commonly encountered than the indifferent words. In our work with younger subjects we have attempted to make such differences minimal. If any residual familiarity differences remain, of a magnitude sufficient to affect the learning process among children, we would expect them to be less important in an older and more highly selected group. The close similarity of results actually obtained for children and for college students may be taken as indicating that in this material the factor of familiarity is not a disturbing element.

A second consideration concerns the universality of emotional connotations of the words. The college students are an older and more sophisticated group of subjects. The finding that similar trends in learning apply to the two different groups is evidence of the generality and of the practical significance of the trends.

A third consideration pertains to a factor which has been suggested as having a possible influence upon our earlier results. In the preceding studies, the children who served as subjects in the learning tests also furnished the evidence as to the affective value of the words learned. The learning tests were given in one laboratory, ratings of the affective values of the words were secured on a different day, and in a different laboratory. The words were also utilized as stimuli to free association in an experiment in which galvanometric and other records of physiological responses were obtained. In some experiments the ratings, and physiological records, preceded the learning tests by about a week, in other experiments these time relationships were reversed. Thus, apparently adequate controls were exercised with reference to the effect of these procedures which were external to the learning experiment. The present data, however, furnish a neater demonstration of the point. The subjects in the present experiment did not rate the words, nor take part in any procedures in which the words were considered as emotional stimuli. Therefore, it cannot be urged that there is any connection between the learning results and the ratings or free associations.

SUMMARY AND CONCLUSIONS

Individual learning tests were administered to 100 college students, in order to check upon certain trends which had previously been found in studies of adolescent children. The tests required the recall of pleasant, indifferent, and unpleasant words in response to pictures which served as paired associates. The data justify the following conclusions.

1. The pleasant words were best learned, the unpleasant words were intermediate, and the indifferent words were learned with least efficiency. The smallest difference is that between *P* and *U*, here the critical ratio is 2.55. The other differences are statistically reliable (critical ratios of 6.04 and 8.31).

2. The trends here found in the learning of college students are the same as those shown in studies of adolescents. The findings earlier demonstrated with the children are evidently not phenomena associated exclusively with a particular developmental period. It is probably justifiable to conclude that the tendencies which have been demonstrated apply to a wide sampling of persons in our present social environment.

3. There is no functional relationship between the learning results and the procedures involved in classifying the words or in free association to the words. The previously established trends in learning have been confirmed with a new group of subjects who were utilized only in the learning experiment and not in the procedures for evaluating the stimuli.

REFERENCES

1. CARTER, H. D., & JONES, H. E. A study of emotional factors in learning and retention. *Psychol. Bull.*, 1933, **30**, 586. (abstract)
2. CARTER, H. D., JONES, H. E., & SHOCK, N. W. An experimental study of affective factors in learning. *J. Educ. Psychol.*, 1934, **25**, 203-215.
3. CARTER, H. D. Effects of emotional factors upon recall. *J. Psychol.*, 1935, **1**, 49-59.
4. CASON, H. The learning and retention of pleasant and unpleasant activities. *Arch. Psychol.*, 1932, No. 134. Pp. 96.
5. JONES, H. E. Emotional factors in learning. *J. Gen. Psychol.*, 1929, **2**, 263-272.

6 MELTZER, H The present status of experimental studies on the relationship of feeling to memory *Psychol Rev*, 1930, **37**, 124-139

7 SMITH, W W Experiments on memory and affective tone *Brit J Psychol*, *Gen Sec*, 1920-21, **11**, 236-250

Institute of Child Welfare
University of California
Berkeley, California

A STUDY OF THE RELATIONSHIP BETWEEN DISCIPLINE AND PERSONALITY TRAITS IN LITTLE CHILDREN*

Pennsylvania State College

MARY ELLEN AYER AND ROBERT G BERNREUTER

In working with adults who are highly unstable it is commonly pointed out that as children they were handled poorly It was thought that such instability would accumulate as a gradual piling up of small insecurities beginning in early childhood. This study is an effort to find out if there is any relationship between the kind of discipline that is used in the home and the personality characteristics of the children who are being disciplined.

THE METHOD USED

Forty children of preschool age, 18 in attendance at the Pennsylvania State College Nursery School and 22 in an Emergency Nursery School in State College, were chosen as subjects of the study The children in the College Nursery School came from the homes of the merchant and professional class; the Emergency Nursery School group was made up of children from homes on the minimum of subsistence level.

The Merrill-Palmer Scale of Mental Tests, administered to both groups, showed that the mental ranking of the children in the College Nursery School was between the 73rd and the 99th percentile, with the median point at the 93rd percentile, and that of the Emergency Nursery School group is between a low 20th percentile and the 99th percentile, with the median point at the 60th percentile

Four aspects of personality, Sociability with Other Children, Attractiveness of Personality, Tendency to Face Reality, and Independence of Adult Affection or Attention, were measured by selecting these four scales from the Merrill-Palmer Personality Rating Sheets Each child was given personality ratings by at least three

---

*Received in the Editorial Office on September 10, 1936

and in some cases as many as seven raters In the College Nursery School, ratings were done for each child by the head teacher and her assistant The other raters in the College Nursery School group were college seniors who had had part-time practice in the school for a semester The group of Emergency School children were rated by three teachers who had daily contacts with the children

An interview for determining the kind and amount of discipline used was devised and used with parents of the preschool children Discipline was divided into these eight fields for the purpose of measuring it.

1 Physical punishment
2. Isolating or ignoring the child
3 Natural result of child's act
4 Worry (scaring child and making him afraid and worried)
5 Rewards or promised rewards.
6 Doing the first thing that pops into a parent's head
7 Temper (on part of parent to get child to do what is wanted)
8. Penance (such as making child sit on chair or go to bed)

Mothers, in every case, and fathers when at home were interviewed The parents were asked if they used a given method of discipline, to give an example of it as used with the child, and to name the frequency with which it was used Frequency of use was divided into five groups· "never," "hardly ever," "now and then," "quite often," "all the time"

These statements of the parents regarding the amount of discipline used were given arbitrary score values from 0 to 4, "0" being assigned to the "never" group and "4" to the "all the time" In order to determine the relationship between the personality rating scores and the method of discipline used, biserial coefficients of correlation were computed An arbitrary dividing line to be used in determining those correlations was drawn between the "hardly ever" and "now and then" groupings This grouped the "never" and the "hardly ever" into one category, and the other statements into the second category

The critical ratios both for the frequency with which various methods of discipline are used in the two nursery schools and for the personality trait scores in the separate nursery schools were obtained by dividing the differences by the standard errors of the differences.

RESULTS

The means and standard deviations of the personality rating scores for each of the two schools is shown in Table 1. This table makes it apparent that the College Nursery School ratings are higher than those of the Emergency Nursery School. From an inspection of the standard deviations shown in Table 1, it is also apparent that the Emergency Nursery School children are more variable than the College Nursery School children. The critical ratios of the differences between the means shows that the difference between the two schools in sociability with other children and tendency to face reality are statistically significant and that the difference in attractiveness of personality approaches statistical significance. The critical ratios of the differences in the standard deviations are significant for all of the traits except independence of adult affection or attention.

TABLE 1

MEANS AND STANDARD DEVIATIONS FOR PERSONALITY RATINGS OF COLLEGE NURSERY, EMERGENCY NURSERY SCHOOL, AND CRITICAL RATIOS OF BOTH COMBINED

	College Nursery School		Emergency Nursery School		Critical ratios	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Sociability with other children	70.8	10.0	61.1	17.8	2.2	—2.5
Attractiveness of personality	62.7	8.7	54.2	21.2	1.7	—3.6
Tendency to face reality	68.8	12.4	55.7	22.0	2.4	—2.5
Independent of adult affection or attention	64.7	12.8	62.2	15.2	0.6	—0.8

The means and critical ratios of the frequency with which various methods of discipline were used in the College and Emergency Nursery Schools are shown in Table 2. The greater use of penance, temper, and the first thing that pops into a parent's head in the Emergency Nursery School are shown by the critical ratios to be significant differences. The score values, showing the more frequent use of worry and physical punishment and the less frequent use of natural result by the Emergency Nursery School group as compared

TABLE 2

MEANS AND CRITICAL RATIOS OF FREQUENCY WITH WHICH VARIOUS METHODS OF DISCIPLINE WERE USED IN COLLEGE AND EMERGENCY NURSERY SCHOOLS

Method of discipline	College mean	Emergency mean	Critical ratio
Physical	1.78	3.18	−1.36
Isolate	2.33	2.00	0.81
Natural result	3.22	2.82	1.61
Worry	0.33	1.46	−1.78
Rewards	2.22	2.27	−0.16
Pop into head	0.78	2.36	−5.29
Temper	1.39	2.32	−3.12
Penance	0.61	2.55	−5.42

with the frequency of use in the other group, approach statistical significance. In this table a minus critical ratio indicates that the method of discipline is used more frequently in the Emergency Nursery School than in the College Nursery School.

The correlations between discipline and personality traits are shown in Table 3. Plus values indicate that the more frequently a

TABLE 3

CORRELATIONS BETWEEN DISCIPLINE AND PERSONALITY TRAITS

Discipline	Sociability with other children	Attractiveness of personality	Tendency to face reality	Independence of adult affection or attention
Physical	+.05	−.17	−.35	−.20
Isolate	−.28	−.06	−.07	+.01
Natural result	+.19	+.33	+.33	+.37
Worry	−.07	−.21	−.30	−.44
Rewards	+.30	−.11	−.14	−.15
Pop into head	−.40	−.54	−.52	−.40
Temper	−.33	−.42	−.61	−.10
Penance	−.15	−.19	−.44	−.08

method of discipline is used the more desirable is the personality characteristic. Minus values indicate that the more frequently a method of discipline is used the less desirable is the personality characteristic.

When physical punishment is used more, there is less tendency to face reality and more dependence upon adult affection and attention. When isolation is used as a method of discipline, the study shows that children have a tendency to be less sociable with other children Natural result used as a means of discipline shows itself to be a good method, for there is a definite positive relationship with it and attractiveness of personality, tendency to face reality, and independence of adult affection or attention As worry is used, scores for attractiveness of personality, tendency to face reality, and independence of adult affection or attention become lower When children receive rewards for doing what they are supposed to do, they tend to have a higher rating in sociability with other children There is a definite negative correlation between doing the first thing that pops into one's head and each of the personality rating factors Temper, used as a disciplinary method, is shown by this study to make children less sociable, to have less attractive personalities, and to face reality less. When penance is used as a method of discipline, children tend to face reality less.

Summary and Conclusions

1 The foregoing paper deals with an attempt to discover the relationship between discipline and personality traits in little children

2 Forty children of preschool age were studied—18 were from the Pennsylvania State College Nursery School and the other 22 from an Emergency School located in State College, Pennsylvania

3. The Merrill-Palmer Personality Rating Scale was used as a method of getting a personality scoring for the children

4. Nursery school teachers and students having frequent contacts with the children were selected as those to rate the children on the personality rating scale

5. Through an interview with at least one parent of each child, the kind and amount of discipline used was determined

6 Biserial coefficients of correlation were computed to determine the relationship between the personality rating scores and the kind and amount of discipline used

7 In general, children in the College Nursery School have a higher personality rating in sociability with other children and in tendency to face reality than do the children in the Emergency Nursery School.

8 Physical punishment, temper, penance, and the first thing that

pops into the parent's head were used much more frequently by the parents in the Emergency Nursery School than in the College Nursery School

9. Methods of isolation, physical punishment, and natural results were used most frequently by the parents taking both nursery schools combined.

10. As physical punishment is used more with children, the children tend to face reality less and show more dependence upon adult affection or attention

11. Attractiveness of personality, tendency to face reality, and independence of adult affection or attention are fostered when children are allowed to profit by the natural results of their acts

12 Worry makes children more dependent upon adult affection or attention, less apt to face reality, and the children disciplined this way have less attractive personalities

13. Giving children rewards for doing things they are supposed to do develops sociability with other children

14 Frequency with which an adult does the first thing that pops into his head has a definite bearing on the personality of the child It makes children have less attractive personalities, be less apt to face reality, be less sociable with other children, and more dependent upon adult affection or attention

15 Using temper in disciplining children makes children less likely to face reality, less sociable, and makes them have less attractive personalities.

16 Penance used as a punishment makes children less likely to face reality.

Pennsylvania State College
State College, Pennsylvania

REMINISCENCE IN MAZE LEARNING BY KINDERGARTEN CHILDREN*

Wesleyan University

GRACE O McGEOCH

PROBLEM

Reminiscence is the improvement in the recall of incompletely learned material after an interval during which there is no formal relearning or rehearsal. A critical review of the experimental literature on reminiscence (3) has shown that many types of verbal materials, e g , poetry, prose, nonsense syllables, lists of words, etc , are susceptible to reminiscence, and that probably the phenomenon is much more prevalent than data on retention in terms of the average score of the group reveal. It was shown that group averages are *inadequate measures* of reminiscence, while the percentages of subjects exhibiting it and the amount shown by them are adequate measures Attention was called to the fact that acts of skill were not in the list of materials which had been shown to be susceptible to reminiscence Decisive evidence of the presence of reminiscence in motor learning would have important bearing on the question of the effect of review during the interval upon the phenomenon

That reminiscence is not necessarily a function of the adventitious factor of review was concluded by McGeoch (3) on the basis of data from two experiments on the learning of poetry by third- and fourth-grade school children This conclusion rested on a comparison of the reminiscence exhibited by the Review Group, which was by far the larger group, and by the No-Review Group A more decisive technique of studying the effect of review upon reminiscence would be the isolation of review by preventing its occurrence In acts of skill such as maze learning the chances of rehearsal during the interval are much smaller than with verbal material Bunch and Magdsick (1) reduced the chances of rehearsal occurring to the minimum by using white rats as subjects in a maze problem

---

*Received in the Editorial Office on September 10, 1936

They concluded on the basis of statistically reliable data that there is an initial rise instead of a decline in the curve of retention under conditions of no intervening practice or rehearsal. Their experimental set-up, however, differed in several respects from that in which reminiscence with verbal materials has been studied. Seven groups, one control group and six experimental groups, were equated in degree of incomplete learning by being given six mechanically guided trials through a multiple-T water maze. The control group, immediately after the guidance, learned the maze to complete mastery (three errorless runs in succession), while the experimental groups did so after intervals of 1, 3, 6, 12, 24, and 48 hours, respectively. Reminiscence was studied by comparisons of the control group with the experimental groups in terms of the average scores of the group (time, trials, and errors) in the final mastery of the maze. The authors state (p. 407) that "the extent to which the partial learning is utilized in the later mastery of the problem, as indicated by the scores in final mastery after the time intervals, is taken as a measure of retention." This is not retention in the sense in which it has been used in previous studies of reminiscence: a comparison of the control group with the experimental groups in performance on the seventh trial would have yielded a measure of retention analogous to the delayed recall in experiments with verbal materials.

There is a suggestion that reminiscence might be present in the results obtained by McGinnis (7) on maze learning by preschool children in that there is a greater decrease in errors from the fifth

TABLE 1
MEAN ERRORS
[Data from McGinnis (7, p. 262).]

	Group I 5 years	Group II 4 years	Group III 3 years
Trial 4	4.25	8.30	7.85
Trial 5	3.85	6.70	6.80
Difference	.40	1.60	1.05
Trial 5	3.85	6.70	6.80
Trial 6	1.85	4.80	5.20
Difference	2.00	1.90	1.60

to the sixth trials, between which there was a 24-hour interval, than from the fourth to the fifth trials which were in immediate succession The data are presented in Table 1 It will be seen from a study of the table that this decrease in errors is greatest with the five-year-old group Since no measures of variability are published, however, the reliability of these differences cannot be determined Furthermore, since there were no control groups, i e, groups who were given six trials in succession, there is no way of determining whether the greater improvement on the sixth trial is actually reminiscence appearing after the 24-hour interval or simply the result of continued practice.

The present experiment was designed to answer this question. Two groups of five-year-old kindergarten children were given five trials in the Young *Slot Maze A* The Control Group took five trials in immediate succession, while Experimental Group I took four consecutive trials and the fifth after an interval of 24 hours.[1] If the children in the two groups are matched in the learning of the first four trials, differences between the two groups on the fifth trial can be considered a function of the 24-hour interval. If reminiscence is present, Experimental Group I should be superior to the Control Group on the fifth trial

EXPERIMENT I · INTERVAL BETWEEN FOURTH AND FIFTH TRIALS

Subjects The subjects were 174 five-year-old children (age range, 60-72 months) of two public school kindergartens in St Louis [2] There were 93 children in the Control Group, 45 boys and 48 girls, and 81 children in Experimental Group I, 47 boys and 34 girls. All the children were of average or above average intelligence according to the results of either the Stanford Revision of the Binet-Simon Intelligence Test or the Pintner-Cunningham Primary Mental Test They were homogeneous in socio-economic status, coming from a very good residential district of St Louis The matching of the children from the Control and Experimental Groups on the basis of performance in the first four trials will be described later

---

[1] *It was planned at first to have the Control Group run six trials in succession,* but preliminary work revealed the fact that six trials made the sitting too long for the slower learners

[2] Acknowledgment is here made to Mrs Rossi and Mrs C Vaulx of the Flynn Park School, and to Miss Gosch of the Jackson Park School for their generous cooperation

TABLE 2
Pairs Matched in First Four Trials

	Time in seconds ($N = 35$)											
	Trial 1			Trial 2			Trial 3			Trial 4		
Group	Mean	σ_M	Range	Mean	σ_M	Range	Mean	σ_M	Range	Mean	σ_M	Range
Control	94.8	15.9	31-534	50.8	15.6	16-403	35.2	5.9	16-181	23.3	7	17-33
Exper I	94.9	15.9	25-523	50.4	10.4	17-322	36.9	6.5	16-207	21.7	1.2	13-40
Diff	—1			.4			—1.7			1.6		
σ_{diff}	22.4			18.7			8.7			1.4		
	Errors ($N = 32$)											
Control	18.9	2.5	1- 77	11.6	2.7	0- 55	4.7	1.5	1- 52	3.2	4	0- 8
Exper I	20.1	3.1	0- 69	10.3	2.3	0- 57	5.3	1.6	1- 51	2.9	4	0-11
Diff	—1.2			1.3			— 6			3		
σ_{diff}	3.9			3.5			2.2			5		

mos ($\sigma_M \pm$ 49) and the mean IQ 115 53 ($\sigma_M \pm$1 96)[6] In Table 2 are given for the 35 pairs the mean time in seconds, the sigma of the mean, and the range for Trials 1, 2, 3, and 4 It will be seen from this table that the greatest difference between the two groups in the mean time for any trial was 1 7 sec (±8 7)

The following facts constitute further evidence of the adequacy of the matching The difference between the pairs was less than 15 sec in 30 of the 35 pairs in Trial 1 and in 32 pairs in Trial 2, while the difference was less than 10 sec in 29 pairs in Trial 3 and in 34 pairs in Trial 4 Such close matching means not only a great similarity between the two groups in the time for each trial but *also in the general course of the learning* with 23 pairs there was a progressive decrease in time with each trial, with four, five, and three pairs there were inversions to increased time in Trials 2, 3, and 4, respectively

The 35 pairs are sufficiently equated in time taken on each of the first four trials to justify treating a comparison of the two groups in the time taken on the fifth trial as a test of the presence of reminiscence. The mean time on the fifth trial, the sigma of the mean,[7] and the range for each group are presented in Table 3

---

[6] It will be noticed that the slight superiority of the Control Group in IQ is not reliable Furthermore, the fact that the correlations between IQ and the various measures of learning, including total time and total errors, ranged from only —.218 (± 083) to + 109 (± 107) indicates that there is little, if any, relationship between IQ and maze learning in the sampling of children used, and justifies the practical necessity of not matching in IQ as well as in time and error scores

[7] These sigmas of matched groups were computed by the following formula given by Lindquist (2, p 199), $\sigma_M = \frac{\sigma\sqrt{1-r^2}}{\sqrt{N}}$ Here $\sigma_{M'}$ represents the sigma of the mean of a matched group, σ, the sigma of the distribution, r, the Pearson product-moment coefficient of correlation between the measures used as a basis for matching and the measures for which the mean is computed In this case the total of the first four trials was taken for the basis of matching The figures given below show that the correlations obtained were all positive but low, and therefore they changed very little, if any, the sigmas of the means computed by the usual formula which are also given

		Time			Errors	
Group	r	P E	σ_M	r	P E	σ_M
Control	+ 11	± 11	1 15	+ 09	±.11	37
Exper I	+ 23	± 11	1 00	+ 20	± 11	.54

TABLE 3

ANALYSIS OF THE FIFTH TRIAL

				Time in seconds ($N = 35$)						
				Improvement			Forgetting			Same score
	Trial 5			Percentage	Amount		Percentage	Amount		Percentage
Group	Mean	σ_M*	Range	showing	Mean	Range	showing	Mean	Range	showing
Control	22 60	1 15	12-44	62 9	4 55	1-11	31 4	6 91	1-23	5 7
Exper I	20 20	97	12-40	60 0	4 95	1-10	31 4	5 00	1-21	8 6
Diff	2 40				—40					
σ_{diff}	1 50				84					
				Errors ($N = 32$)						
Control	3 07	37	0-10	46 9	2 06	1- 4	31.2	2 80	1- 7	21 9
Exper I	3 31	53	0-19	50 0	1 93	1- 5	31.2	4 20	1-14	18 7
Diff	—24				.13					
σ_{diff}	65				39					

*Sigma of mean of matched groups

Study of this table will show that it took Experimental Group I 2.40 sec. less time, on the average, than the Control Group. This difference, however, has a sigma of 1.50. Furthermore, an analysis of the two groups on the basis of the percentage of subjects showing improvement and forgetting and the mean amount shown revealed no significant differences between the Control Group and Experimental Group I. The results are also given in Table 3. Examination of these results reveal the fact that the two groups are remarkably similar on the fifth trial: 62.9 per cent of the children in the Control Group compared to 60.0 per cent of Experimental Group I improved, i.e., took less time on the fifth trial than on the fourth; 31.4 per cent in each group forgot, i.e., took longer on the fifth trial than on the fourth, while 5.7 per cent of the Control Group compared to 8.6 per cent of Experimental Group I took the same time on the fifth as on the fourth. Furthermore, the difference of .40 sec. in favor of the Control Group in the mean amount of improvement shown is not significant since its sigma[8] is .84.

These results in terms of time scores yield no evidence of the presence of reminiscence in the fifth trial. The conclusion that reminiscence is not present is corroborated by the following analysis of the error scores.

Pairs Matched in Errors. Thirty-two pairs of subjects were found who were equated in errors on the first four trials. The mean CA of the 32 children from the Control Group was 66.40 mos. ($\sigma_M \pm .55$) and the mean IQ was 118.28 ($\sigma_M \pm 2.31$). The 32 children from Experimental Group I had a mean CA of 68.06 mos. ($\sigma_M \pm .66$) and a mean IQ of 113.46 ($\sigma_M \pm 2.58$). The mean number of errors with its sigma and the range in error scores of each group for the four trials are given in Table 2. The greatest difference between the two groups in the mean number of errors for any one trial was 1.3 errors (± 3.5) in Trial 2. The closeness of matching is indicated further by the following facts. In Trials 1 and 2, 23 and 27 pairs, respectively, differed by less than five errors, in Trial 3, 30 pairs differed by less than four errors, while in Trial 4 all pairs differed by less than four errors. Again there was great similarity between the pairs in the general course of the learning, with 28 pairs there was a progressive decrease in errors with each trial, while there was an inversion in the second trial with three pairs and in the third trial with one pair

---

[8] It is interesting to note here also the mean amount of forgetting exhibited by the 11 subjects in each group. The small numbers preclude statistical treatment, but it is highly doubtful if there are significant differences between the two groups in this respect.

Again we look to the fifth trial for evidence of reminiscence. The relevant data are found in Table 3 Study of this table shows that there is practically no difference between the Control Group and Experimental Group I in the mean number of errors in the *fifth trial, since the difference of — 24 has a sigma of 65* Likewise in terms of the percentage of subjects who exhibited fewer errors, more errors, or the same number of errors on the fifth trial compared to the fourth the two groups are very similar 46.9 per cent of the Control Group compared to 50 0 per cent of Experimental Group I improved, 31 2 per cent of each group forgot; and 21 9 per cent of the control Group compared to 18.7 per cent of Experimental Group I had the same error score Furthermore, *there is no significant difference between the two groups in the mean* amount of improvement, since the difference of 13 has a sigma[9] of 39. Thus the results in terms of error scores, as those in terms of time scores, do not yield evidence of reminiscence in the fifth trial

The explanation of these results may lie in the fact that the phenomenon of reminiscence is a function of the material and is not present in the learning of the Young *Slot Maze A* by five-year-old *children, or in the fact that reminiscence is conditioned* by other factors, such as the length of the time interval, which in our experimental set-up were not favorable to its appearance Degree of learning or stage in the learning process might be another such factor.[10] It is possible that there is a stage of learning for the insertion of the interval of time optimal for yielding reminiscence. Perhaps after the fourth trial it was too late in the learning process to produce reminiscence. That the Control Group had reached a negatively accelerated stage of learning can be seen from the data given in Tables 2 and 3.

---

[9] As in the analysis of the time scores, the number of subjects showing forgetting is too small to treat statistically It is, nevertheless, interesting to note that, with the omission of the one child in Experimental Group I who made 14 more errors on the fifth trial than on the fourth, the mean number of errors for the remaining nine children is 3 11 This is only 31 above the mean of the 10 subjects from the Control Group.

[10] The experimental literature on verbal learning permits no definitive statement regarding the effect of degree of learning upon reminiscence (5)

EXPERIMENT II: INTERVAL BETWEEN SECOND AND THIRD TRIALS

To test the possibility that reminiscence might occur at an earlier stage in learning, another group, which will be designated Experimental Group II, was used with whom the 24-hour interval was inserted between the second and third trials. The procedure in all other respects was the same as that used with the Control Group and Experimental Group I. The 64 subjects in Experimental Group II, 35 boys and 29 girls, were five-year-old kindergarten children from the same schools and, therefore, similar to the other subjects in socio-economic status and intelligence.[11]

RESULTS

If reminiscence is present, it should be apparent in a comparison of the Control Group with Experimental Group II on the third trial, provided the two groups have been equated in learning in Trials 1 and 2.

Pairs Matched in Time. There were 35 children in the Control Group (mean CA 67.54 mos. $\sigma_M \pm .52$) who could be paired with 35 children in Experimental Group II (mean CA 66.25 mos. $\sigma_M \pm .64$) in the time taken for the first and second trials. With 6 of the 35 pairs the difference in time in either trial was between 11 and 15 sec., with 8 pairs, between 6 and 10 sec., with 21 pairs, 5 sec. or less. The closeness of the matching is further evident from a study of the data given in Table 4 which include the mean time and its sigma[12] and the range of the two groups for each trial.

---

[11] Intelligence test scores were not available for Experimental Group II, but, as was discussed before, there is little relationship between intelligence scores and maze learning with the sampling of children used. Furthermore, intelligence has not been demonstrated to be a factor in reminiscence.

[12] These sigmas of means of matched groups were computed by the formula given earlier in this paper. The total of the first two trials was the basis

TABLE 4

Pairs Matched in First Two Trials

Group	Trial 1 Mean	Trial 1 σ_M	Trial 1 Range	Trial 2 Mean	Trial 2 σ_M	Trial 2 Range	Trial 3 Mean	Trial 3 σ_M*	Trial 3 Range
Time in seconds ($N = 35$)									
Control	128.3	19.9	31-485	49.9	6.9	18-238	43.31	6.44	12-170
Exper II	128.1	19.7	29-491	51.6	7.4	19-228	34.28	2.98	13-118
Diff							9.03		
σ_{diff}							7.09		
Errors ($N = 45$)									
Control	20.9	2.7	2- 90	11.9	1.7	1- 47	12.71	3.00	0- 82
Exper II	20.7	2.8	3- 88	11.6	1.8	1- 46	7.64	1.32	0- 39
Diff							5.07		
σ_{diff}							3.27		

*Sigma of mean of matched groups

Turning to Trial 3 for evidence of reminiscence, we find a difference of 9.03 sec in mean time in favor of Experimental Group II. This difference is not reliable, however, since its sigma is 7.09. Further analysis of the third trial yielded no evidence of reminiscence. 71.4 per cent of each group showed improvement, and there was no significant difference between the two groups in the mean amount of improvement shown (Control Group, 27.6 sec. ±8.6, Experimental Group, II 31.2 sec. ±6.9).

This failure to find reminiscence in the time scores was corroborated by the following results in terms of errors.

Pairs Matched in Errors. Forty-five children in the Control Group (mean CA 66.62 mos $\sigma_M \pm .47$) were matched with 45 children in Experimental Group II (mean CA 66.48 mos $\sigma_M \pm .49$). The closeness of the matching is indicated by the data given in Table 4. The following facts constitute further evidence, with 34 of the 45 pairs the difference in either trial was two errors or less, while with the remaining 11 pairs the difference was between three and five errors.

Again turning to the third trial for evidence of reminiscence we find the difference of 5.07 errors in favor of Experimental Group II has a sigma of 3.27 and is, therefore, not significant. Furthermore, although 64.4 per cent (±7.2 per cent) of the Control Group compared to 51.1 per cent (±7.5 per cent) of Experimental Group II showed improvement, the difference here of 13.3 per cent has a sigma of 10.4 per cent. Likewise there is no reliable difference between the two groups in the mean amount of improvement shown: Control Group 9.1 (±2.0) errors, Experimental Group II 12.5 (±2.1) errors.

DISCUSSION

Under our experimental conditions reminiscence is not present in maze learning by kindergarten children. It is highly doubtful, how-

---

of matching. The correlations and the sigmas of the means, computed by the usual formula, are as follows:

Group	Time r	 PE	 σ_M	Errors r	 PE	 σ_M
Control	+.19	±.11	6.58	+.22	±.09	3.10
Exper. II	+.53	±.08	3.55	+.39	±.09	1.43

ever, that this failure to find reminiscence should be considered to imply that acts of skill in general are not susceptible to reminiscence. Attention has been called elsewhere (3) to the relationship between reminiscence and the *superiority of distributed practice in learning*, a phenomenon occurring in the learning of acts of skill as well as of verbal materials. The optimal stage in the course of learning for the insertion of a rest interval is a function of many conditions (6). It is highly probable that conditions determining reminiscence have a differential effect upon the individual members of a group in any experimental set-up.[13] Therefore, a rest interval inserted at an arbitrarily chosen stage in the learning of a group, e.g., after two trials or four trials, would not yield reminiscence in all members of that group. This would be particularly true of a maze pattern of great unevenness in difficulty. The Young *Slot Maze A* is an unfortunate selection because of the fact that in one section of it (alley 9) behavior commonly described as "insight" occurred: the subject caught on to the fact that he had to move away from the goal, which was spatially then very near, in order to reach it.[14] *The fact that the data show remarkably little effect of any kind of* the insertion of the 24-hour interval upon the learning[15] suggests that the effects of the interruption of practice were less important than the particular stage of the learning, i.e., the presence of "insight" with respect to this most difficult section of the maze. If the 24-hour interval had been inserted at the *critical* time for each individual with respect to his catching on to the fact (and his accompanying emotional state) that he must move away from the goal at the entrance to Alley 9, there might have been evidence of reminiscence The factors determining this *critical* time constitute, of course, a part of our main problem, the conditions of reminiscence

---

[13] Results (4, pp. 107-108) on the learning of two sets of materials under identical conditions by the same subjects did not yield evidence of the constancy of reminiscence in a given learner, i.e., some subjects were not reminiscers and others non-reminiscers. This suggests that the phenomenon is determined more by the conditions of learning than by the learners.

[14] McGinnis (7, p. 260) also found with her five-year-old group the greatest number of errors in this section.

[15] The great similarity on the fifth trial between the Control Group and Experimental Group II in terms of reminiscence and forgetting has been discussed above. *It might be argued, however, that, although the effects of* the interruption of practice might not be apparent on the first trial after the 24-hour interval, subsequent trials might reveal some effects. With Experimental Group II the data on the fourth trial as well as the third trial were available. Analysis of the fourth trial gave the same results; in time or error scores, there were no consistent or reliable differences between the Control Group and Experimental Group II.

In the beginning of this paper the importance of the presence of reminiscence in motor learning with respect to the question of the effect of review during the interval upon the phenomenon was mentioned The most decisive technique of studying the effect of review upon reminiscence is the isolation of review by preventing its occurrence In acts of skill, such as maze learning, and with very *young human subjects or animals, the chances of unintentional rehearsal* during the interval are reduced to a minimum The suggestion in the data of McGinnis (7) that reminiscence is present in the learning of the Young *Slot Maze A* by kindergarten children has not materialized The problem should be attacked with a more uniformly difficult act of skill

SUMMARY AND CONCLUSIONS

The purpose of this experiment was to discover whether reminiscence occurred in maze learning with kindergarten children The Young *Slot Maze A* was used with 241 five-year-old children of average and superior intelligence, who were divided into three groups a Control Group who took five trials in succession, Experimental Groups I and II with whom an interval of 24 hours was inserted between the fourth and fifth trials, and the second and third trials, respectively Children in the Control Group were matched in time scores and error scores, on the first four trials with those in Experimental Group I, and on the first two trials with those in Experimental Group II On any one trial the majority of the members of a pair did not differ in time scores more than 10 sec and in error scores more than two errors Comparisons were made, in terms of time and errors, of the fifth trial of the Control Group with the fifth trial of Experimental Group I, and of the third trial of the Control Group with the third trial of Experimental Group II There is no consistent or reliable evidence of reminiscence in either Experimental Group

REFERENCES

1 Bunch, M E, & Magdsick, W K The retention in rats of an incompletely learned maze solution for short intervals of time *J Comp. Psychol*, 1933, **16**, 385-409

2 Lindquist, E F The significance of a difference between "matched" groups. *J Educ Psychol*, 1931, **22**, 197-204

3 McGeoch, G O The conditions of reminiscence *Amer J Psychol*, 1935, **47**, 65-89

4 ——— The age factor in reminiscence a comparative study of preschool children and college students *J Genet Psychol*, 1935, **47**, 98-120

5 ——— The factor of degree of learning in reminiscence a second comparative study of preschool children and college students *J Genet Psychol*, 1935, **46**, 455-462

6. McGeoch, J A Learning Chap 13 (esp pp 320-322) in *Psychology a factual textbook*, ed by E G Boring, H S Langfeld, and H P Weld New York Wiley, 1935

7 McGinnis, E *Motor habits in young children* *Genet Psychol Monog*, 1929, **6**, 203-311

Wesleyan University
Middletown, Connecticut

A FURTHER STUDY OF EARLY CHILDHOOD MEMORY*

Ohio State University

HAROLD E. BURTT

The present study is a continuation of one reported earlier (1) in which nonsense material, presented systematically to the subject in infancy, was relearned at a later age and the rate of relearning compared with the rate of learning new material of similar nature The previously reported relearning experiment began at the age of 8½, the present one at the age of 14. The material consisted of selections from Sophocles' *Oedipus Tyrannus* in the original Greek, which was tantamount to nonsense material Each selection included approximately 20 lines or 240 syllables of iambic hexameter The subject was a boy with an IQ of approximately 130 At the age of 15 mos, three of these selections were read to him once daily for a period of 3 mos—a total of 90 repetitions At the age of 18 mos these selections were discontinued and three others read daily for 3 mos This procedure was continued until the subject was three years old and 21 selections had been presented The 8½-year experiment utilized only one-third (seven selections) of the available material which had been presented in infancy plus three new selections. The 14-year experiment employed another third of this original material plus three other new selections, and makes possible a comparison of the "saving" at these two ages

The procedure at 14 years was practically identical with that at 8½, except that the distribution of trials was a trifle more uniform Two trials were given daily for the most part One interruption of five days occurred at about the 150th trial after several of the selections were already completed and out of the way The entire 10 selections under comparison were always given in a trial, and their order from trial to trial was rotated systematically For instance, if one trial was 2, 5, 8, 11, 14, 17, 20, *D, E, F*, and next would be 5, 8, 11, 14, 17, 20, *D, E F* 2 As in the earlier ex-

---

*Received in the Editorial Office on September 20, 1936

periment, the 10 selections were merely read to the subject for the first 18 trials Beginning with the 19th, every third trial used the prompting method in which the selection was read slowly and the subject anticipated any words or syllables which he could These words were underlined in the text and the date noted at which they had first been anticipated correctly This procedure was continued until every word in the selection, with the exception of the initial cue words, had been anticipated and the requisite number of trials constituted one item of score The repetitions were continued if necessary until the subject could recite the entire selection without prompting, and the number of trials up to this point constituted the second item of score The attitude of the subject in the present case was somewhat superior to that in the earlier experiment because he understood the scientific importance to a greater extent Motivation seemed entirely adequate throughout Toward the end there was a little unavoidable further motivation as certain selections were learned and dropped from further consideration and effort was made to get still others out of the way

A variable which could not be controlled was the occasional tendency for the subject to think of words or phrases between trials A little beyond 100 trials he began to recite phrases spontaneously when riding, walking, or swimming According to his report at the end of the experiment, he seldom did this except when in the company of the experimenter On this basis the occasions were not very numerous The experimenter made no comment on such occasions After trial 118 the subject remarked that some of the selections were nearly learned, and quoted part of selection *XX* to illustrate his point. A few days later he realized that *XX* was complete except for one group of words and he tried to quote it between trials in order to determine where he was "stuck" Subsequently, he listened intently to that part of the selection The above is the only instance that the experimenter observed in which the conduct of the subject gave any selection an undue advantage It is probable that this consideration between trials affected the selections about equally When questioned at the conclusion of the experiment, the subject reported that no selections had been particularly impressive and he had no notion as to which ones were being relearned and which ones learned *de novo*

The main results, together with some data from the 8½-year experiment, by way of comparison, appear in Table 1 The first column gives the age at which the original reading took place The Roman numerals in the second column indicate the arbitrary designation of the selections that were used in the 8½-year experiment. The two following columns give the number of repetitions necessary before the subject had anticipated each word in the selection and also the repetitions required in order to recite it verbatim without prompting. The remaining columns give similar data for the 14-year experiment The two selections involved in a row of the table had been presented in infancy under identical conditions The averages for certain groups of selections are given in the rows indicated The lower portion of the table gives the data for selections which were learned *de novo*—*A*, *B*, and *C* for the 8½-year experiment and *D*, *E*, and *F* for the 14-year experiment

TABLE 1

Age in mos at original reading	Eight-year experiment			Fourteen-year experiment		
		Repetitions			Repetitions	
	Selection	Anticipate each word	Recite entire	Selection	Anticipate each word	Recite entire
15-18	III	361	382	II	142	142
18-21	VI	250	253	V	133	139
21-24	IX	376	385	VIII	148	169
24-27	XII	355	379	XI	148	151
27-30	XV	304	328	XIV	145	145
30-33	XVIII	217	226	XVII	169	169
33-36	XXI	259	265	XX	127	127
Av of all		303	317		145	149
Av 3 early		329	340		141	150
Av 4 later		284	299		147	148
Control	*A*	409	409	*D*	163	169
	B	445	451	*E*	145	151
	C	436	445	*F*	163	166
Av control		430	435		157	162

The most general notion of the results may be obtained from the averages It requires 145 repetitions before each word or syllable in the selection is anticipated in the relearning experiment, while with entirely new selections of comparable difficulty 157 repetitions are necessary The corresponding averages, when it is a matter of actually reciting the selection without prompting, are 149 and 162

A considerable overlapping of the distributions is apparent, as some of the new selections require less repetition than some of those which are relearned The saving averages about 8 per cent As a small number of items contribute to each average, it is necessary to analyze the difference between averages according to the small sample theory (Fisher) · In both of the above cases the probability, by this method of computation, is about 0 8 that the difference is real Thus, from a statistical standpoint the differences cannot be considered entirely significant. It is interesting to compare the present experiment with the earlier one In that case the saving was about 30 per cent as contrasted with the present 8 per cent, and the differences were of undoubted statistical significance. The intervening six years produced a considerable loss in retention for the material presented in infancy

Continuing the analysis in the same way that it was made with the first set of data, the relearning is divided into two portions—Selections II, V, and VIII which were originally presented when the subject was youngest, and Selections XI, XIV, XVII, and XX which were given later in infancy The averages appear in the rows of the table labelled "Av. three early" and "Av four later," respectively. In the 8½-year experiment the selections presented later in infancy actually required less repetition for relearning on the average. In the present instance this tendency has disappeared Again, the seven figures in a given column may be ranked in order of magnitude and those ranks correlated with the order in which the selections were presented originally in order to discover any tendency for the later selections to be learned more readily than the earlier ones In the 8½-year experiment a correlation of 50 was obtained in this way. In the present case the correlation is zero

Learning curves are shown in Figure 1. The abscissa indicates the trial and the ordinate the cumulative number of syllables correctly anticipated up to and inclusive of the trial The solid curves are for 8½ years and the dotted ones for 14 The upper curve in each *instance is for relearning and the lower for learning de novo* The separation of the curves was scrutinized for its statistical significance Typical ordinates were taken at intervals of 21 trials and the significance of the difference between the ordinates for two curves computed in accordance with small sample theory In the 14-year curves,

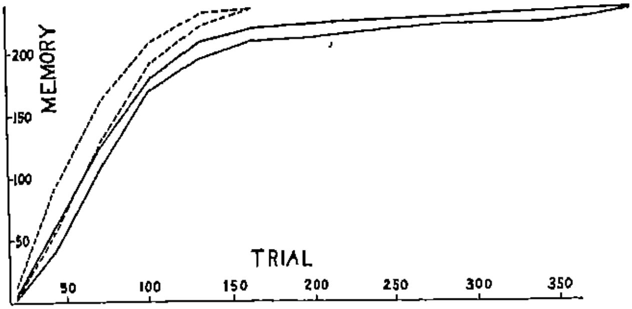


FIGURE 1

after the first few trials, these differences have probabilities greater than 99 that the difference is real. This condition exists until near the 80th trial when the probability becomes approximately 98 Thereafter the significance of the differences decreases as the curves approach one another. In the above respect the present experiment differs from the earlier. In that case, practically none of the differences between ordinates was significant The result is perhaps attributable to the slightly greater homogeneity of data in the present case. It appears that the selections came along together to a greater extent than they did in the original experiment.

Another difference between the present and the earlier study is the much smaller number of repetitions necessary for both relearning and learning For example, the three new selections required 162 repetitions for a complete recital in the present case, whereas in the earlier study they required 435. This difference is due partially to the greater maturity of the subject and to better motivation. He appreciated the scientific importance of the experiment and was also aware of the fact that when he had learned a selection he would be through with it The main interest of the study, however, was in the comparative saving rather than the absolute scores

On the whole, the experiment suggests that traces of the effect of presenting nonsense material in infancy were present at age 14 but that the effect was unquestionably less marked than at age 8½. We still have available seven more selections which were presented in infancy and with which a third experiment will be possible at some future time.

SUMMARY

Meaningless material (20-line selections of Greek drama) was read aloud to the subject daily, beginning at the age of 15 mos. Every 3 mos a different set of similar selections was used as material, and this procedure was continued until the age of three. When the subject was 8½-years old he learned some of the original material by a modified prompting method in comparison with other material which he learned *de novo*. A similar experiment was conducted with still other material from the original selections at the age of 14

At the age of 8½ the relearning required approximately 30 per cent fewer repetitions than were needed to learn new material. At the age of 14 the corresponding figure was 8 per cent. At the age of 8½ the original selections which had been presented later in infancy were relearned more readily than those presented earlier in infancy By the age of 14 this difference had disappeared Learning curves, however, show some significant differences between the ordinates for learning and relearning at the age of 14 The effect of the presentation in infancy was very marked at the age of 8½, but at the age of 14 the effect, while still appreciable, was considerably decreased

REFERENCE

1 BURTT, H E An experimental study of early childhood memory *J Genet. Psychol*, 1932, 40, 287-295

Ohio State University
Columbus, Ohio

SHORT ARTICLES AND NOTES

THE EFFECT OF PENALTY UPON FREQUENCY OF STUTTERING SPASMS*

C. VAN RIPER

One of the features with which the psychologist must constantly deal when investigating stuttering is the great amount of variation in the frequency of stuttering blocks or spasms. Steer (2) has shown that frequency of stuttering spasms is, at least to a certain degree, a function of the type of speech situation. Van Riper and Hull (4) showed a progressive decrease in frequency of spasms as the stutterer adapted to any given speech situation. However, little knowledge was gained from these experiments as to what factors in the situations were responsible for the variations in frequency.

It was this experimenter's hypothesis that in any given speech situation one of the factors which determines the frequency of stuttering spasms is the felt or expected penalty which attaches to the stuttering spasm.

In order to determine the exact rôle played by the expected penalty, the following procedure was devised and administered to 16 stutterers. The stutterer was seated in the same room with two experimenters. Electrodes connected with a Harvard induction coil were fastened to his neck. He was then given a reading passage which included all the sounds of the alphabet in their initial positions, and asked to read the passage six times, a one-minute rest being given between readings. During the pause before the fourth reading, the stutterer was told that, for each spasm which occurred during that reading, he would be given a severe electric shock, the shocks to be given at the conclusion of that reading. A sample shock was administered. *The two experimenters, independently of each other,* recorded the number of spasms per reading by underlining the words stuttered upon. It was found that this threat produced an average increase of 5.2±0.8 over the preceding reading, all but one of the subjects showing an increased amount of stuttering. This is of greater significance, in view of the finding of Van Riper and Hull that the number of spasms normally tends to decrease from reading to re-reading. The added penalty *placed upon stuttering, therefore, produced more spasms,* a finding of immediate importance to clinicians.

In order to insure, however, that it was fear of a penalized stuttering rather than fear of electric shock which was responsible for this increase,

---

*Received in the Editorial Office on May 19, 1936.

two further readings were used as controls. No shock was threatened or given on the fifth reading, but prior to the sixth reading the subject was told that he would be given as many shocks as he had had spasms on the initial reading, but that the number of spasms which he had would in no wise affect this number of shocks to be given at the end of the reading. This threat was found to produce an average increase of only 1.5±0.7 spasms over the previous reading, approximately one-third of the subjects showing no increase whatsoever.

When computed according to the usual methods, it was found that there were 99 chances out of 100 that threat of shock per spasm would produce more stuttering than threat of shock regardless of spasms, the average difference being 4.5±0.6 in favor of the former, all but one of the subjects showing a greater amount of stuttering in this direction. In order to eliminate errors due to sequence, one-half of the subjects were given the threat of shock alone on the fourth reading and threat of shock per spasm on the sixth reading. No essential differences were found whether the data for the two experimenters were averaged or treated separately. The data included in Table 1 represent the average of the two experimenter's independent counts.

TABLE 1

DIFFERENCES IN FREQUENCY OF SPASMS BETWEEN READINGS

Compared readings	Average diff	S D
Reading 1—Reading 2	3.9	3.6
Reading 2—Reading 3	3.3	3.1
Preceding reading—Threat of shock reading	—1.5	2.7
Preceding reading—Threat of shock per spasm reading	—5.2	3.0
Threat of shock per spasm reading—Threat of shock reading	4.5	2.6

It is felt that the results of the reported experiment show conclusively that frequency of stuttering spasms is in part, at least, a function of the penalty felt by the stutterer to be attached to them.

The importance of such a finding to therapy is obvious. If one is to decrease the number of stuttering spasms, it is necessary to decrease the penalty attached thereto. Fundamentally, what society penalizes is the interruption to communication and the abnormal manner in which communication is finally achieved. Since the only symptoms common to all stutterers are the repetitions and prolongations of articulatory postures, and since the variation in symptoms from stutterer to stutterer seems largely to be a function of the habitual reactions he makes to these primary symptoms (3), it would seem advisable to teach the stutterer a manner of stuttering which minimizes the interruption and abnormality as much *as possible*. Such a therapy is at present being experimentally developed at the State University of Iowa Speech Clinic. Similarly, this experiment points toward the development of mental-hygiene techniques in the direction of those advocated by Bryngelson (1).

From a theoretical point of view, the results of this experiment bear on the problems concerning differences in the amount of stuttering in the so-called "easy" and "hard" situations, and upon the great individual differences among stutterers.

REFERENCES

1. BRYNGELSON, B. A method of stuttering. *J. Abn. & Soc. Psychol.*, 1935, 30, 194-198.
2. STEER, M. D. Influence of social and psychological factors upon severity of stuttering. M.A. Thesis, State Univ. of Iowa, 1933.
3. VAN RIPER, C. The systematic treatment of stuttering. *Proc. Amer. Speech Correc. Soc.*, 1936, 110-120.
4. VAN RIPER, C., & HULL, C. J. The quantitative measurement of the effect of certain situations on stuttering. (In press.)

State University of Iowa
Iowa City, Iowa

---

*A NOTE ON THE PERSEVERATING TENDENCY IN STUTTERERS**

JON EISENSON

We are interested in learning whether stuttering is a form of perseveration. Perseveration is the tendency of ideas to remount into consciousness spontaneously, after they have once occurred. H. M. Cushing (1) defines perseveration as a lag or inertia of chemical-neural processes.

To date, investigators have been concerned chiefly with the phenomenon of perseveration in the abnormal. Ewen (2) investigated perseveration in normal and epileptic subjects. He found the perseverating tendency to be a feature of the peculiar mental state of epileptics between attacks. H. H. Jasper (3) *defines perseveration as the tendency* of a set of neurons, once excited, to persist in this state of excitation autonomously, showing resistance to any change in this state. J. W. Pinard (4) found that perseverators tend to be nervous and sensitive. The reoccurrence of a tune or a jingle or the feeling of being on roller skates after the skates have been removed are common examples of the perseverating tendency.

---

*Received in the Editorial Office on May 19, 1936.

This paper will deal with the results of a preliminary investigation of the perseverating tendency in stutterers. Is the tendency of the stutterer to repeat a sound that he has made a manifestation of the perseverating tendency? The production of a speech sound results from the excitation of a set of neurons. If the state of excitation continues autonomously, resisting change, the result will be a repetition of the original sound produced.

In order to determine the existence of a general perseverating tendency among stutterers, a series of tests[1] was administered to a group of 30 stutterers and a control group of 30 non-stutterers. The groups were matched for chronological age, mental age, and school environment. The age range was from 10 to 16. Both groups were composed entirely of males. The test material was constructed along these lines: (1) A situation is presented to which the subject is required to respond for a given period of time. (2) A variation of this situation is then presented to which the subject is required to make a different response for the same time period. (3) A third situation, consisting of the combination in random order of both original situations, is then presented. The subject is here required to change his reactions quickly and suddenly. Failure to make these changes is a failure in responding to a new situation. A failure means that the subject continued to react to a situation which was no longer present. A mental process continued in activity after the conditions to which it was originally due were no longer present. This is taken as a measure of the perseverating tendency.

Another factor entering into the measure is the number of items completed within the allotted time. The fewer the items completed in changing from one task to another the greater the amount of time needed for the adjustment to the new task, or, arguing conversely, the greater the amount of time spent per item the larger the degree of perseveration.

The complete battery of 15 tests included the following types of mental and motor tasks: (*a*) writing letters, (*b*) drawing horizontal and vertical lines, and (*c*) addition and multiplication of one-place numbers.

The results of the test are summarized in Table 1. The stutterers completed *more items than the non-stutterers in those situations which called* for a continuation of the same reaction to an unchanging stimulus (280 items for stutterers compared with 262.7 for the non-stutterers). The difference between the means of the groups is 17.9 items. This difference approaches statistical reliability $\left(\dfrac{\text{Diff}}{\sigma_{\text{diff}}} = 1.2\right)$. A larger number of cases

---

[1] These tests were part of a battery included in the Maller-Elkin Attention Test for the Measurement of Perseveration, Teachers College, Columbia University, 1935.

TABLE 1

COMPARISON OF STUTTERERS AND NON-STUTTERERS IN TESTS OF PERSEVERATION

	Stutterers	Non-Stutterers	$\frac{\text{Diff}}{\sigma_{diff}}$
Mean perseveration quotient	63.5	67.7	2.7
Mean initial score	280.6	262.7	1.2
Mean final score	524.7	527.3	1
Mean errors	7.0	4.2	1.4

would be needed, however, to establish unquestioned statistical significance.

Despite the advantage in number of items completed in the initial tests, the non-stutterers *completed more items in the complete test battery than* the stutterers (527.3 for non-stutterers, 524.7 for stutterers). The mean number of errors for the stutterers was 7.03, for the non-stutterers 4.17. The ratio $\frac{\text{Diff}}{\sigma_{diff}}$ is 1.4.

There are two things to be noted. First, the stutterers "slowed-down" more in their tasks than the non-stutterers, and, second, the stutterers made more errors while working. The slowing down and the errors we consider to arise out of the stutterer's difficulty in making ready adjustments to a changing situation. He failed to respond to new stimuli as accurately and as rapidly as the non-stutterer.

A perserveration quotient based upon the initial and final scores was computed for each of the groups. According to the method of computation, the larger the quotient the smaller the degree of perseveration. The mean P. Q. for the stutterers was 63.5, for the non-stutterers, 67.7. The difference between the two groups (4.2) is statistically significant. The ratio $\frac{\text{Diff}}{\sigma_{diff}}$ was 2.72.

SUMMARY AND CONCLUSIONS

The results of the tests show a greater tendency on the part of the stutterers to resist change—a greater tendency for their neurons, once excited, to persist in the original state of excitation than is the case with normal speakers.

We propose the theory that stuttering itself is an indication of resistance to change, and hence a manifestation of the phenomenon of perseveration.

REFERENCES

1 CUSHING, H M. A perseverative tendency in pre-school children *Arch. Psychol*, 1929, No 108

2 EWEN, J H Perseveration in the insane epileptics *J Ment Sci*, 1930, **76**, 537-540.

3 JASPER, H H Is perseveration a functional unit participating in all behavior processes? *J Soc Psychol*, 1931, **2**, 28-51

4 PINARD, J. W Tests of perseveration *Brit J Psychol*, 1932, **23**, 114-126

Brooklyn College
Brooklyn, New York

INCIDENCE OF ORPHANHOOD AMONG FIFTEEN HUNDRED PSYCHOTIC PATIENTS*[1]

H BARRY, JR AND W A. BOUSFIELD

The number of writers who stress parental influence as a factor in normal development is impressive Campbell (6), Hartwell (11), Myerson (19), Henderson and Gillespie (13), Thom (24), and Wallin (25), to cite only a few, attest the value of an adequate child-parent relationship. In view of the findings of these authorities, it might be expected that the disrupting of the home by the death of parents would be followed by some noticeable change in the personality make-up of the child It is also obvious that the various ramifications of this problem comprise a subject for intensive study and observation. In the field of criminology this is true to the extent that broken homes have been reported as a causal factor in delinquency by a long series of investigators including Healy and Bronner (12), Glueck and Glueck (9), Breckinridge and Abbott (3), Shideler (21), Sullenger (23), and Erickson (8) In contrast to the number of the above researches, there is relatively scant information concerning this approach to the genesis of emotional disorders. Plant (20) briefly presents disintegrated family relationships as a psychiatric problem. Grzywo-Dabrowska (10) refers to the large percentage of orphans among a series of juvenile suicides, and Barry (2) reports on the high incidence of orphanhood among psychotic rulers

The present investigation deals with 1500 psychotic cases, all of them white, and first admissions to the New Jersey State Hospital, Greystone

---

*Received in the Editorial Office on September 14, 1936

[1] The writers wish to acknowledge their indebtedness to Drs Marcus Curry, Superintendent, and A. G Lane, Clinical Director, of the New Jersey State Hospital, Greystone Park, New Jersey

Park, New Jersey. Data regarding family background were obtained from the anamneses comprising part of the routine clinical case histories. In this report we have adopted Lotka's (16) definition of an orphan as a person who has lost one or both parents by a certain specified age. The arbitrary age of 12 was selected since it appeared adequate to include those cases for whom informants reported, "The mother (father) died while the patient was a child." Furthermore, this age serves as a convenient approximation of the onset of puberty.

The data indicate that the patients born prior to 1905 include a high percentage for whom adequate information could not be obtained. While there might be a presumption that these unknown backgrounds would include a relatively large proportion of orphans, this is too uncertain to serve as a basis for any final conclusion. For the cases born subsequent to 1905, however, the records are reasonably complete, as indicated in Table 1.

TABLE 1

INCIDENCE OF ORPHANS AMONG FIFTEEN HUNDRED PSYCHOTIC PATIENTS

Family situation at the age of 12 (percentage of total cases)

Date of birth	Total cases	Orphaned	Not orphaned	Information inadequate
1916 *et seq*	53	26.4	71.7	1.9
1911-1915	186	28.0	66.1	5.9
1906-1910	249	18.9	77.1	4.0
1901-1905	234	20.5	65.4	14.1
1891-1900	457	18.2	62.8	19.0
1881-1890	321	19.6	59.8	20.6

There is a marked increase in the incidence of orphanhood for the patients born subsequent *to 1910. Combining the data for this group of younger* inmates, we find that 27.6 per cent lost one or both parents before they were 13 years old. This is 8.7 per cent higher than the figure obtained for the slightly older group born between 1906 and 1910, in spite of the fact that mortality rates and consequent expectation of orphanhood have decreased steadily since the beginning of the century [cf. Lotka (15)]. Comparison between these subgroups taken from a single hospital population strongly suggests that there is a more pronounced relation between orphanhood and the subsequent development of a psychosis in the cases of patients who are admitted before they are 25 years of age.

Aside from the relative differences between the figures for the younger and older patients, the actual magnitude of the incidence of orphans in the younger group is impressive. While it would appear improbable that even 20 per cent of the general population would lose a parent by the age of 12, control figures are none-the-less needed. Published data on this subject are relatively meager. The results compiled in Table 2 are admittedly inadequate as a basis for comparison, but they do furnish an indication of the normal expectancy of orphanhood. It may be seen that the figures presented in Table 2 are consistently below those presented in Table 1.

TABLE 2

INCIDENCE OF ORPHANS AMONG NORMAL INDIVIDUALS

Investigator	Per cent orphans	Comment
Burdge (4)	15.0	New York state: ages 16-18
Burt (5)	17.0	London, 1921: influenced by war casualties
Lotka (16)	14.8	British census, 1921: influenced by war casualties
Lotka (16)	16.7	Computed for U. S. based on mortality tables
Slawson (22)	15.9	New York City school children: grades 5-10

It should be noted that other investigators have presented statistics on broken homes, including in this category divorce, desertion, and separation. Typical of these studies are those of Baker, Decker, and Hill (1), Campbell (7), Monroe (18), and Mead (17). Their percentages range from 18 to 23. While these figures indicate a fair degree of consistency with each other, and with the data for orphans previously quoted, it should be evident that they are not strictly comparable, since they represent varying ages at which the broken home factor is considered. Moreover, as Slawson (22) has indicated, the percentage of broken homes varies widely between different districts. He presents results for three different schools which vary from 12.3 per cent to 24.1 per cent according to the social status of the neighborhood. Obviously more research must be undertaken before the reliability of these results is definitely established.

CONCLUSIONS

1. Family backgrounds were studied for 1500 psychotic patients. Of those born since 1910, 27.6 per cent had lost one or both parents by the age of 12.
2. Comparable figures for normal controls as presented by different investigators are definitely below 20 per cent.
3. The evidence suggests that the older patients have a smaller incidence of bereavements than those who develop a psychosis before the age of 25.
4. Results presented here are preliminary until confirmed with a larger group.

REFERENCES

1 BAKER, H J, DECKER, F J, & HILL, A S A study of juvenile theft. *J Educ. Res*, 1929, **20**, 81-87.

2 BARRY, H, JR Orphanhood as a factor in psychoses *J Abn & Soc Psychol*, 1936, **30**, 431-438

3 BRECKINRIDGE, S P., & ABBOTT, E The delinquent child and the home New York Russell Sage Foundation, 1912

4 BURDGE, H G Our boys Albany State of New York Military Training Commission, 1921 [Cited by Slawson (22)]

5 BURT, C The young delinquent New York Appleton, 1925

6 CAMPBELL, C M. Human personality and the environment New York Macmillan, 1934

7 CAMPBELL, M W The effect of the broken home upon the child in school *J Educ Soc*, 1932, **5**, 274-281

8 ERICKSON, M H Some aspects of abandonment, feeblemindedness, and crime *Amer J Soc*, 1931, **36**, 758-769

9 GLUECK, S, & GLUECK, E T One thousand juvenile delinquents Cambridge Harvard Univ Press, 1934

10 GRZYWO-DABROWSKA, M Samobojstwa mlodziezy szkolnejna podstawie ankiety (An inquiry on suicide in high schools) *Oswiata i Wychowanie*, 1932, **4**, 38-75, 133-163 (Cited from *Soc Sci Abst*, 1932, **4**, 1337)

11 HARTWELL, S W Adult adjustments and non-adjustments in relation to their effects upon children *Ment Hygiene*, 1932, **16**, 598-609

12 HEALY, W, & BRONNER, A F Delinquents and criminals, their making and unmaking New York Macmillan, 1928

13 HENDERSON, D K, & GILLESPIE, R D A textbook of psychiatry for students and practitioners London Oxford Univ Press, 1933.

14 KAWIN, E Children of pre-school age Chicago Univ Chicago Press, 1934.

15. LOTKA, A. J. The diminishing burden of orphanhood a great social benefit *Stat Bull, Metropolitan Life*, 1933, **14**, 1-5

16 ——— Orphanhood in relation to demographic factors *Metron*, 1931, **9**, 37-109

17 MEAD, L L Report on the survey of truancy for the year 1929-1930 *Commonwealth Rev* (Univ Oregon), 1931, **13**, 245-265

18 MONROE, D. Chicago families, a study of unpublished census data Chicago Univ Chicago Press, 1932 [Cited in Kawin (14)]

19 MYERSON, A. The foundations of personality Boston Little, Brown, 1921

20 PLANT, J S Sociological factors challenging the practice of psychiatry in a metropolitan district *Amer J Psychiat*, 1929, **8**, 705-716

21 SHIDELER, E H Family disintegration and the delinquent boy in the United States *J Crim Law & Criminol*, 1918, **8**, 709-732

22 SLAWSON, J The delinquent boy Boston Badger, 1926

23 SULLENGER, T E Social determinants in juvenile delinquency (Ph D Thesis, University of Missouri, 1929)

24 THOM, D. A. Everyday problems of the everyday child. New York: Appleton-Century, 1934.

25 WALLIN, J. E. W. Personality maladjustments and mental hygiene. New York: McGraw-Hill, 1935.

Tufts College
Massachusetts

APPARATUS

A SAND-TUBE OBSTRUCTION APPARATUS*

Calvin P. Stone

The accompanying figure illustrates an apparatus unit that was designed by the author to make use of the rat's digging proclivities in overcoming an obstruction placed between himself and a specific desideratum. The unit illustrated in this figure is being used to study maternal drive. With slight alteration it can be used for a variety of other problems in the field of motivation.

Cage *A* is the incentive chamber. In this during the training trials is placed an object that is appropriate to a strong internal need in the animal being tested. For example, one places here an inverted tube of water for the thirsty, a pan of food for the hungry, nesting materials or infant young for the mother, a receptive female for the adult male, an adult male for the sexually receptive female, etc.

Tube *B*, extending from Cage *C* to *A*, is the receptacle for the sand obstruction. This tube consists of a piece of sheet-metal tubing[1] 24 in. long and 3 in. in diameter, to which at the lower end a 90-degree elbow of the same material is attached. The free end of the elbow is so trimmed that its open mouth fits into the elliptical hole in the floor of Cage *C* and projects above the floor approximately ¼ in. A 1-in. strip of hardware

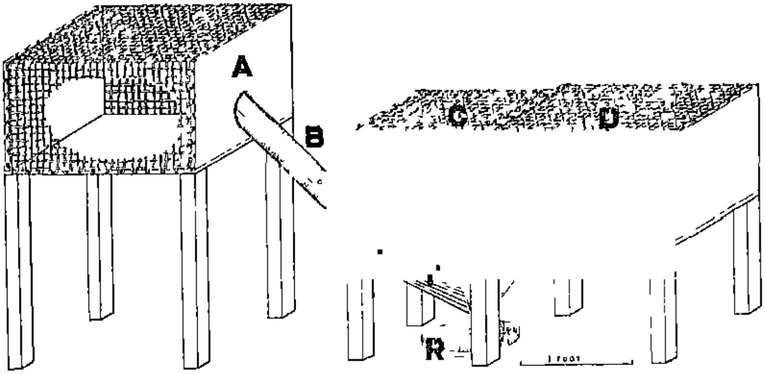


FIGURE 1

---

*Received in the Editorial Office on May 19, 1936.

[1] A piece of spouting or stone pipe.

cloth fastened lengthwise in the tube enables the rat to ascend and descend the tube with ease. At the upper end of the tube a trap-door is placed. This may be kept open altogether, or opened by the animal as it ascends the tube. In the latter case, the door closes behind the animal and thus prevents his returning down the tube. When another animal is confined in the incentive chamber as a lure, the closed trap-door prevents the animated lure from entering the tube. Fine sand is poured into the upper end of the tube by means of a long funnel. *As a rule, the rolling sand* completely fills the lower portion of the tube, but if this is not the case a small amount of sand should be added at the distal end by the experimenter. One varies the total amount of sand according to the amount of work one wishes to impose upon the animal. A volume that stands 8 or 10 in. in the tube can be removed in a minute or less by a strongly motivated rat that has had a few successful experiences in removing this obstruction.

Cage *C* is the digging chamber. Its sides and floor are made of hardware cloth. This makes the animal visible to the experimenter and allows the sand to disappear as it is pawed from the tube. A metal trough (*T*) beneath the floor of Cage *C* collects the sand as it falls through the floor and assembles it in a receptacle (*R*). In most experimental settings the animal is started in Cage *C*, but in studies of maternal behavior the home cage of the mother (Cage *D*) is jutted against the end of Cage *C* and an opening between these cages is made to allow the mother freedom of movement to and from the home cage. Cage *D* is a permanent part of the present unit. In it the mother and her young are placed a day or two before critical trials on maternal drive are made. When several animals are to be tested, portable maternity cages that can be set upon a small stand of the appropriate height are substituted for Cage *D*. This procedure reduces the time required for preliminary adaptation of the mother rat to the test situation.

SUGGESTIVE USES OF THE SAND-TUBE OBSTRUCTION APPARATUS

This apparatus has been used at Stanford University in undergraduate courses for the following purposes:

1. *Correlation of Deprivation Interval with Animal Drive.* By way of preliminary training, a rat is allowed to explore all parts of the unit thoroughly, with no sand whatever in the tube. Then food, for example, is placed in the incentive chamber and the animal is allowed to make a few trips to the food and to eat a few pellets at the end of each trip. Next, about half of the lumen of the tube is filled with sand. Under this condition no digging is required, but the rat will have to crawl through a smaller opening. Following this, a trial is given with the lumen approxi-

mately three-fourths closed Then comes a trial with the lumen completely filled Finally, a few more trials are given with the lumen completely filled in order to stabilize the animal's reactions to the obstruction This completes the preparatory training for the experiment proper in which the deprivation interval is varied after the manner of Warner (1) who used the Columbia Obstruction Apparatus Rats are tested for the number of instances they will remove the obstruction during a standard test period of 20 min These data which may be obtained in a short time by a group of students working cooperatively make possible a rough, but satisfactory, co relation of the efforts of the animal with the interval of food deprivation In a similar manner, tests of thirsty animals can be made

2 *Insight Reactions* A hungry rat that has been allowed to explore all parts of the apparatus is permitted to obtain food in the incentive chamber From time to time he is returned by the experimenter to Cage *C* in order to force him to ascend the tube to reach food. So far the tube is entirely empty When the hungry animal has acquired the habit of ascending the tube immediately upon being dropped into the digging chamber, the lumen of the tube is filled with sand, and the behavior of the animal as he encounters and finally removes the obstruction is observed By the use of a code and a mimeographed sheet, with which the student has familiarized himself in advance, a comprehensive record of the animal's activities in the digging chamber is charted on a time line If the animal fails in 20 min, he is removed and given other trials on the same or the following day until he solves the problem Of special interest is the contrast between the rat's activities when he first encounters the obstruction and his behavior in trials after he has mastered the obstruction This is a very satisfactory experiment for beginning students because the animal almost always solves the problem in the first or second critical trial.

3 *Studies in Perseverance* Taking animals that have had experience in removing the sand obstruction and are consistent from trial to trial, one may perform studies on the perseverance by suddenly increasing the amount of work to be done The latter is accomplished by inserting a 1-in. rubber tube into the upper end of the sand-tube and pouring sand through a funnel into the tube as the sand is removed by the rat Prolonged periods of work in which the animal's behavior differs markedly from the accustomed behavior of one that has been successful in overcoming the obstruction within a minute or two are thus evoked Repeated experiments of this kind afford a point of departure for discussions of the literature on animal frustration

4 *Individual Differences* By using a number of animals in each of the foregoing experiments, data are assembled for the study of individual differences In this connection, differences that can be ascribed to varia-

tions in techniques are readily differentiated from those arising when techniques of the experiment are held constant

Up to the present time no attempt has been made to lay down standard methods for the use of the sand-tube obstruction apparatus in research It is assumed that each experimenter will wish to develop his methods in accordance with the nature of his problems

Reference

1 Warner, L H A study of hunger behavior in the white rat by means of the obstruction method *J Comp. Psychol*, 1928, 8, 273-299

Stanford University
California

*A PAPER-WINDOW OBSTRUCTION APPARATUS**

Calvin P Stone

Haggerty used a paper window in studies of imitation in monkeys (1) This usage suggested to the author the possibility of employing the basic idea of the paper window in obstruction tests with rats Some of the ways it has been employed by the author in studies of motivation and learning will be described

In Figure 1 are given three views of a typical window frame constructed to hold a piece of paper toweling in place This frame consists of two pieces of three-ply wood, 6 in. wide and 12 in. long, hinged together at the top so that the two pieces may be separated and brought together in sandwich fashion At the ends of these boards opposite the hinge, 4-in holes have been cut Near the edge of the hole in one of the two boards short nails were driven from the outer to the inner face so that their free ends project into holes in the inner face of the opposite board These nails, four or five in number, hold a piece of paper lying between the sandwich boards firmly in place when the boards are pressed together as illustrated in Figure 1*B* Figure 1*C* illustrates a frame with paper that has been torn by a rat preparatory to going through the opening

Frames of different dimensions have been used by the author and also openings of different shapes and areas have been tried Apparently a considerable latitude of variation in these details is permissible without materially altering the usefulness of the paper as an obstruction The circular openings shown in Figure 1 were cut by means of a fly-cutter which oper-

---

*Received in the Editorial Office on July 23, 1936

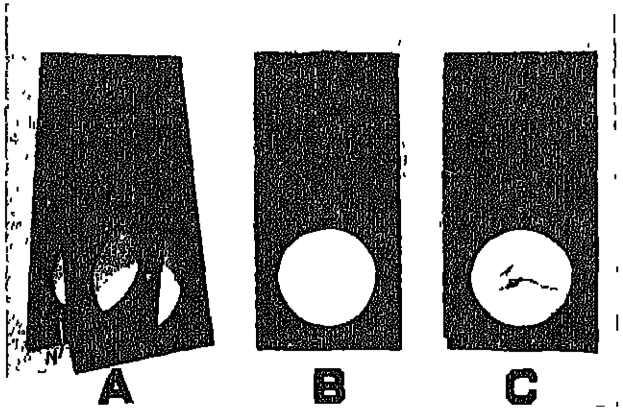


FIGURE 1

THREE VIEWS OF THE WINDOW FRAME USED IN THE PAPER OBSTRUCTION TEST They are 6 in wide and 12 in long. Element A shows the frames ready for the insertion of a piece of paper toweling (the obstruction) which is held in place by small nails (*n*) when the inner faces of the two boards are pressed together, as in *B*. Element *C* illustrates the type of hole made by a rat, preparatory to going through to reach the food.

ates in an ordinary machinist's lathe. By this means a perfect circular opening of any desired diameter may be cut in approximately one minute.

Figure 2*A* illustrates a demonstration apparatus consisting of three complete units for obstruction tests. The top and outer walls are made of hardware cloth. The window frames stand vertically and are held in place by means of wooden guides nailed on the outer walls and the partitions of the box. The frames may be lifted out for the insertion of new paper but cannot be displaced by the animal.

To use this apparatus a rat is put into one of the front compartments (*S*) and is required to pass the paper barrier by tearing a hole in it to reach the goal compartment (*G*). Ordinarily at the outset some preliminary training is given in order that the animal will be well oriented toward the task of overcoming the barrier. This consists of giving three or four opportunities to pass through the unobstructed opening to reach food in the rear compartment and then two or three additional trials with a piece of torn paper in place such as that in Figure 1*C*. On the last of these trials the animal is required to enlarge the hole in the paper by tearing it with his mouth or by forcing his body through the opening.

A good class demonstration is afforded by confronting three animals with

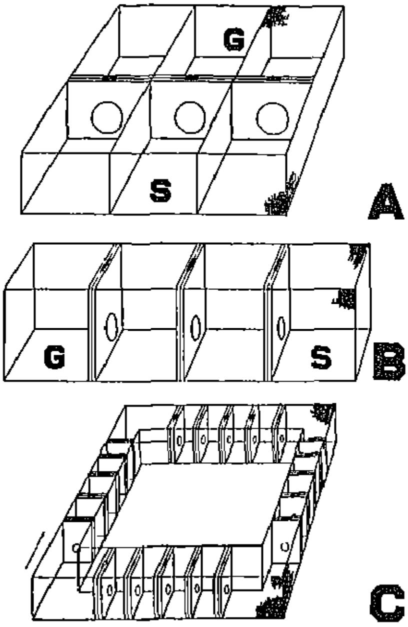


FIGURE 2

ELEMENT *A* ILLUSTRATES A THREE UNIT DEMONSTRATION BOX DESIGNED PRIMARILY FOR CLASSROOM USE

It is 30 in. wide, 24 in long and 12 in high Element *B* is a straight runway accommodating three obstruction frames It is 4 ft long, 12 in. high, and 6 in wide. Element *C* is a quadrangle 6 ft. long, 6 ft wide and 12 in. high. The width of runway is 6 in It accommodates 20 paper obstruction frames

different degrees of training with the paper obstruction test simultaneously In the first compartment is placed a hungry rat that has been allowed to complete the first step of preliminary training in which it merely goes through the unobstructed hole to the food, in the second compartment is placed a hungry rat that has completed the second step in preliminary training, namely, that of going through a large hole in the paper made by the experimenter, and in the third will be put a hungry rat that has completed both steps of the preliminary training and, in addition, on one previous occasion has torn a hole in an intact piece of paper Under these conditions the most experienced animal will immediately rend the paper window as shown in Figure 1*C* and dash through to obtain the reward, the next most experienced animal will center its activities upon the window, but ordinarily will leave it frequently before sinking its teeth into the intact paper, an act that is the first indication of a solution of the problem, the least experienced animal will display a great deal of desultory behavior and seldom make an opening in the paper window during one observation period of 15 min., unless given aid, such as a small perforation in the paper made by running a lead pencil through it.

Figure 2*B* illustrates an apparatus that was designed to accommodate three windows in a straight runway. With this the animal may be required to break through three barriers in succession in order to pass from the start (*S*) to the *goal* (*G*). The object of this arrangement of barriers is to provide way stations for studying the behavior of animals toward ultimate and subsidiary goals in motivational studies involving incentives that have different preferential values.

Figure 2*C* illustrates a still more extended series of windows, 20 in number, so arranged that the animal comes back to the starting-place, or to an incentive chamber that is near it (but not shown in the drawing). If one has need of only one straight runway, he uses only one side of the quadrangle and shuts it off from the rest by means of a solid board that replaces one of the window frames. Also, if fewer than the maximal number of frames is desired on either side of the quadrangle in any experiment, those in excess of the present need are removed from the apparatus. This particular adaptation of the basic idea of the paper obstruction was devised to provide a suitable work-load to differentiate between strongly and weakly motivated rats, and also between those that were appropriately rewarded and those that were inappropriately rewarded. The differentiation rests on qualitative differences in behavior and on quantitative differences in the number of barriers overcome in a time-limited observation period.

No doubt other experimentalists will find many diverse ways in which to employ the basic idea of the paper obstruction. The author has illustrated only a few of these possibilities. The paper obstruction is recommended chiefly because it has the characteristics of a genuine barrier and at the same time is one that may be overcome again and again by rats of any age beyond weaning without setting up undesirable fear or avoidance reactions toward it.

REFERENCE

1 HAGGERTY, M. E. Imitation in monkeys. *J. Comp. Neur.*, 1909, **19**, 337-455.

Stanford University
California

BOOKS

HARVEY A CARR *An Introduction to Space Perception* New York Longmans, Green & Co, 1935. Pp xi+413

More than seventy-five years have passed since Lotze, Helmholtz, and Wundt undertook to formulate the problem of space perception. This problem has since become one of the most fruitful and perhaps *the* most important of all the problems in experimental psychology Nevertheless, the problem of space perception is so intertwined with other aspects of sensory function that it has hardly achieved the degree of autonomy that is fully warranted in view of a voluminous experimental literature. Timely, indeed, is this authoritative English text devoted exclusively to these materials

The work evinces a distinctly modern cast One finds here a treatment that is not weighed down by extended historical or encyclopedic documentation The functional approach is consistently maintained, and the problem of efficiency is kept in the focus

The general problem is oriented primarily in respect of *remote* objects (p vi). The author's objectives are to indicate the various modes of apprehending spatial relations, to specify the factors upon which such functions essentially depend, and to set forth the conditions under which they are optimally determined. Representative studies which bear upon the various phases of the problem are summarized and critically evaluated

The book comprises a short introduction (pp 1-6), ten longer chapters (*vide infra*), and subject and author indices (pp 407-413) In two cases, a heading is allocated two chapters, otherwise, the subject-matter of different chapters is essentially distinctive Nevertheless, the successive chapters are bound into a unity by a thread of continuity, to wit, a *motor* (proprioceptive) theory of space perception, which usually is explicated after a presentation of experimental data This theory specifies that spatial relations are apprehended in terms of motor adjustments (manipulations) in respect of objects Adjustment for each point in the spatial field is distinctive and is the basis of a specific *locality reference* ("local sign"). It is in terms of the efficiency with which these adjustments are made that the capacity of the organism for apprehending spatial relations may be determined

That *space* is a geometrical abstraction—a "conceptual" rather than a perceptual matter—is stated at the very outset. The senses act cooperatively in apprehending spatial relations (Chapter 2, pp 7-58) However, vision is said to be primary it is the single receptor-system in terms of which these relations usually are most accurately determined, and it is dominant over hearing in localizing with the aid of pseudophones The cooperative action of the senses in this context is *acquired;* a new functional relation is readily established between vision and somesthesis (proprioception) after the customary one has been disrupted by means of a lens that shifts visual relations in a definite way

Chapter 3 (pp 59-91) is devoted to the inverted retinal image, or optical inversion, mystery It is indicated that the correlative fact of horizontal inversion has been generally ignored The enigma in this much-discussed problem has its basis in a dualistic distinction between a physical and a visual or psychological object, the latter of which has no existence independent of and apart from the perceiving subject

The treatment in Chapters 4 and 5 (pp 92-158) concerns the auditory perception of space. The author here attempts to distinguish between visual and auditory perception of space, i e., in terms of vision spatial relations are *sensed* but in terms of audition they denote *meanings* The latter is said to signify a relationship between an auditory object and a visual and somesthetic object However, the distinction is one of degree rather than one of kind, since visual relations of space denote meanings between a visual object and a visual and somesthetic object

Chapter 6 (pp 159-194) is primarily concerned with the fact that excitations in the two *corresponding areas* in the binocular relation entail like localizing responses and signify a common locality reference, while excitations in *non-corresponding* areas refer to objects at different spatial positions. Chapters 7 and 8 (pp 195-288) are devoted to visual perception of distance, and Chapter 9 (pp 289-335) contains the treatment of visual perception of direction. Knowledge regarding the latter problem is less definite and precise because it is more complex and has not been investigated so thoroughly. In Chapter 10 (pp 336-358) the topics of *continuous* and *discontinuous* movement are summarized within the bounds of a short paragraph apiece, and the topics of relativity in the perception of motion, objective and organic criteria, and after-effects are discussed at some length. The last chapter (pp. 359-405) is concerned with linear and areal aspects of visual perception

A treatment which makes efficiency the central problem has a very definite bearing upon the problem of learning The latter dovetails with the problems of space perception in two distinctive ways (1) the rivaling claims of nativists and empiricists need to be evaluated in order that the scope and the bearing of the problem of learning in the present context may be specified, and (2) the course (character, rate, etc) of improvement in efficiency needs to be traced in connection with such aspects of the problem as manifest the property of acquisition

The first of these phases of the learning problem, that concerning the distinction between native and acquired aspects of space perception, is one of the salient considerations in the present work It receives attention in a number of connections The reader may be a bit surprised to discover that the author leans so strongly towards nativism at several points. This is illustrated in the specification that the common locality reference of corresponding areas in the binocular relation is "innately conditioned." Although the comment is added that the strength of the connection may be materially influenced by experience (p. 180), this commitment has the effect of shifting learning to a rôle of secondary importance.

The second, the *developmental* phase of the learning problem receives treatment by implication at least in such a context as that regarding the establishment of a new functional relation between vision and somesthesis (proprioception) when devices are used to alter the normal spatial relationships. This aspect of the learning problem does not receive the careful attention that it deserves, and possibilities are not seized upon to indicate instances in which developmental features may be revealed presently. Thus, in the last chapter, pronouncement is made that size and shape must be taken for granted as a given datum of experience (p. 361). A different treatment could have indicated that the manner in which these characteristics emerge from non-spatial components has not yet been revealed clearly, but it is quite conceivable that in due time it, too, will be shown to be a function of learning.

The author's decision to restrict references largely to English sources is well justified. Specific reference to the significant contributions of Lotze and Wundt in such a work is a matter of choice and not of requirement. Besides, statement was made in the *Preface* that little use was to be made of the wealth of German literature. However, the neglect of such other German sources as are now readily accessible in English is not so easily comprehended. Too, it cannot merely be taken for granted that the (American) functional and the (German) configurational approaches which converge upon this problem are mutually exclusive. The laws of spatial grouping and other *factual* data are not specifically considered in the very contexts in which they have a natural setting. A more definite comment in the *Preface* could have dispelled much wonder as regards this indifference.

These critical comments are here set forth for the purpose of indicating points at which the reviewer believes the treatment might have been made more representative. They are not to be construed as seriously militating against the end for which the book was intended. The work is, indeed, what it purports to be—an introduction to space perception. For this purpose, it is very suitably adapted. As the only authoritative (modern) text upon the topic of space perception, it embodies points of high merit. Chief among these are the functional approach and the designation of efficiency (the adequacy of the organism in responding to the environment) as the central problem in perception.

Wellesley College
Wellesley, Massachusetts

MICHAEL J. ZIGLER

The Journal of Psychology

EDITED BY

Carl Murchison

The Journal of Psychology was founded at the request of many psychologists throughout the world for the purpose of establishing a medium in which publication shall be immediate and in which other publication experiments may be carried out. Beginning with the third volume the following methods and terms will be in force:

(1). The size of each volume will be increased to 500-600 pages, and the subscription price will be $7 per volume, two volumes per year.

(2). At the request of the American Library Association, the loose leaf feature will be discontinued, and only the regular library edition will be sent to subscribers. The format of each volume will continue to be such that each subscriber may tear down the assembled volume into its component articles.

(3). Except through the fault of the author, the lag in publication should not be more than three or four weeks.

(4). The author will continue to receive 200 reprints of his paper, and the subscriber will receive the accumulated publications bound into half volume units. The second half of each volume will contain the title page and table of contents for that volume. Each half volume unit will be mailed to the subscriber at its stated date.

(5). The restriction on the length of experimental papers is removed, but theoretical papers must not be more than sixteen pages in length.

(6). Each author will be charged the actual manufacturing costs of any zinc or half tone cuts.

(7). The publication charges will be readjusted as follows:

4	page	signature	$16
8	"	"	20
12	"	"	30
16	"	"	40
20	"	"	50
24	"	"	60
28	"	"	70
32	"	"	80
36	"	"	90
40	"	"	100

These charges are tentative and will be adjusted as justified by the subscription list.

(8). *The Editorial Board.* To a much greater extent than is the usual custom the Editorial Board will be drawn from the ranks of the younger psychologists who have shown research promise and have achieved a rank equal to full membership in the American Psychological Association.

(9). The accumulated units will be mailed to subscribers quarterly: January, April, July, and October. The units may vary in size, to accommodate publication needs.

Beginning with January, 1937, one thousand to twelve hundred pages (two volumes) annually. Per annum $14.00; per volume $7.00; per half volume unit $4.00. Complete sets at $7.00 per volume, plus transportation.

THE JOURNAL PRESS
2 Commercial Street
Provincetown, Massachusetts
U. S. A.

$7.00 per volume
Single numbers $4.00

QUARTERLY
Two volumes per year

June, 1937
Volume 50, Second Half

Founded by G. Stanley Hall in 1891

THE PEDAGOGICAL SEMINARY AND
JOURNAL OF GENETIC PSYCHOLOGY

Child Behavior, Animal Behavior, and Comparative Psychology

JUNE, 1937

(Manuscripts are printed in the order of final acceptance)

The development of temperature sensitivity during the fetal period . . . 217
By Leonard Carmichael and G. F. J. Lehner

Serial Learning II. The bi-directional goal gradient in the endless maze . . . 229
By Karl F. Muenzinger, Claude C. Dove, and Arthur H. Bernstone

How rats learn the simple alternation problem in a temporal water maze . . . 243
By Sol Evans

Continued training with the omission of certain nebenreize . . . 277
By Paul R. Farnsworth

A study of the Cornell-Coxe performance ability scale with superior children . . . 283
By Hazel Lincoln

The mechanical abilities of siblings . . . 293
By Ellis K. Frye

The effect of functional periodicity on the learning process . . . 307
By Orpha Maust Lough

Correlations of chronological age, mental age, and intelligence quotient with other abilities and traits in Grade I . . . 323
By Frank T. Wilson and Cecile White Flemming

(OVER)

Copyright, 1937, by The Journal Press
Provincetown, Massachusetts

Entered as second-class matter August 3, 1897, at the post-office at Worcester, Mass., under the act of March 3, 1879

Reentered as second-class matter May 11, 1937, at the post-office at Provincetown, Mass., under the Act of March 3, 1879

Personality ratings. I Reliability of teachers ratings 339
By T A Langlie

Scholastic achievement of Indian pupils 361
By Harvey C Hansen

The development of visual acuity in the cat 371
By J. Warkentin and K U Smith

Personality ratings. I Validity of teachers ratings in predicting scholastic achievement 401
By T A Langlie

The mental development of children from underprivileged homes 427
By Harold M Skeels and Eva A Fillmore

The free word association of elementary school children 441
By William McGehee

SHORT ARTICLES AND NOTES 457
Some characteristics of the written speech of stutterers (Jon Eisenson) 457

Plantar responses in infants following a startle stimulus (Frances M Clarke, William A Hunt, and Edna B Hunt) 458

APPARATUS 463
The typewriter key as a device for registering choice Timing (Marie L H. Forbes) 463

BOOKS 465
N Kohts, Infant ape and human child, instincts, emotions, play, habits (Arnold Gesell) 465

BOOKS RECEIVED 468

THE DEVELOPMENT OF TEMPERATURE SENSITIVITY DURING THE FETAL PERIOD*

University of Rochester and Brown University

LEONARD CARMICHAEL AND G. F. J. LEHNER

I. INTRODUCTION

In 1885 William Preyer wrote in his significant book on the special physiology of the embryo

> It is of physiological and especially of psychogenetic importance to determine experimentally when in man and animals the individual sensory nerves become excitable, how the neonate and the fetus in particular react to tactual, thermal, electrical, and chemical skin stimuli, to gustatory and olfactory stimuli and to auditory and visual stimuli (15)

After more than half a century, this program, the significance of which is even more obvious than it was in Preyer's time, has still not advanced far

At the present time, the only sense which has been at all thoroughly studied in its full development in fetal life in typical mammals has been that of cutaneous pressure, although, of course, incidental observations on the development of a number of other receptor fields, especially including cutaneous pain, have been made on human and mammalian fetal material during this period

The experiments reported in this paper deal with the development of the temperature senses during the behaviorally active period of a typical mammalian fetus The philosopher, John Locke (11), in the seventeenth century attributed temperature sensitivity to the fetus Preyer (15), Kroner (9), Kussmaul (10), Genzmer (7), Blanton (2), Peiper (12), Canestrini (4), Avery (1); Pratt (13); Pratt, Nelson, and Sun (14), Jensen (8); Carmichael (5, 6), and others, have all referred to temperature sensitivity in the fetal or newborn human or mammalian organism So far as the writers are aware, however, the study reported here is the first

---

*Received in the Editorial Office on October 15, 1936

systematic investigation of the effect of temperature stimulation during the whole fetal period in a typical mammal It is also possibly the first study of any receptor field beside that of cutaneous pressure to be studied during this complete period

II. Experimental Procedures

Guinea-pig fetuses of known insemination age were prepared for study, using procedures previously described by one of the writers (6) By the use of these techniques, fetuses of any desired insemination age can be exposed in a thermostatically controlled bath of body-warm physiological salt solution, with placental circulation maintained. Appropriate lights and a motion-picture camera, with special lens equipment, make recording of the fetal behavior possible

Six stimulus temperatures were used in this study Three, namely, 47° *C*, 67° *C*, and 85° *C*, were above the physiological zero of the organism, taken to be 37 5° *C.*, and three 27° *C*, 17° *C*, and 5° *C*, were below A supply of physiological salt solution at each of the temperatures noted above was maintained in separate vacuum flasks, each of which was fitted with an individual pipette and an individual thermometer

After the fetus had been exposed and observed in a preliminary way, it was gently lifted just above the surface of the bath, and three, or in certain instances five, drops of liquid taken from the bath and hence of a temperature closely approximating the physiological zero, were dropped from a pipette held just above the surface of the skin on to certain definite cutaneous areas On each fetus the six spots studied were (*a*) vibrissae-area, (*b*) ear, (*c*) shoulder, (*d*) rump, (*e*) fore-paw, (*f*) hind-paw Following this in various orders, but ordinarily with the extreme temperatures at the end, a similar procedure was employed for each of the six temperatures listed above. After each stimulation, time for recovery was provided while the fetus was again immersed in the bath At the conclusion of this series, the fetus was ordinarily again tested with drops of liquid at the physiological zero temperature This procedure was adopted to determine what responses were released by the tactual component, as contrasted to the behavior elicited by the temperature component of the drops of liquid used as stimuli In using this technique, the assumption is made that responses elicited by drops of liquid of temperatures other than that of the bath and which were not elicited by stimulation of similar drops of liquid at the physiological zero of the organism may be considered as being released by thermal stimuli. The possibility must be remembered that at certain temperatures and at certain ages, pain receptors, not temperature receptors proper, were stimulated by the liquid stimuli of the temperatures used. Strictly speaking, this paper therefore reports a study of responses to thermal stimuli and not directly a study of the so-called temperature end-organs.

In all, 58 fetuses were used in this study. Of these, 18 were eliminated from consideration in the results presented, because they were not in normal condition or because complete observations were not made upon them. The 40 fetuses actually studied came from 21 litters of gestation ages varying from 28 days and 10 hours to 62 days and 2 hours. The first motility of the guinea-pig has been observed by Bridgman and Carmichael (3) at approximately 25 days. The normal gestation period for the guinea-pig is ordinarily set at 68 days.

III. Experimental Findings

One response, a combined shoulder and limb reaction, was observed in one of three fetuses of a litter 28 days and 20 hours old. This response was released by drops of the coldest stimulus, 5° C., applied to the shoulder. No other temperature-released responses were observed in this litter, although tactual stimulation with a hair did release behavior in each case. Temperature stimulation may release behavior in fetuses younger than the one reported above, but it did not do so in this study.

The general results of the experiment are presented in five tables and in the discussion of these tables. It should be remembered that the units counted in making up these tables have in many cases required judgments on the part of the experimenters concerning the nature of the response in question, which could readily have been challenged at the time. However, other qualitative (i.e., the as yet unquantified) aspects of behavior, as observed by the experimenters but not recorded, seemed to confirm the quantitative generalizations given below in a most decided way. Both of the writers were present at every experiment and agreed on the protocols as recorded at the time.

Table 1 gives the number of maternal animals and the age and

TABLE 1

Maternal Animal	Fetuses Studied Age	Fetuses Studied Length in Mm.	Weight in Grams
	Young Group		
A	28 20	22 32 23 6 23 75	1 45 1 45 1 65
B	30.18	26 85	2 08
C	32 19	45 44 5 43 25 40 75	6 8 7 3 6 8 6 4
D	33	50 75	9 79
E	34 18	35 7	4.2
F	35	37 75 38 35	4 35 4 4
G	35 17	45 4 42 7	6 47 6 03
H	36 22	44 7 45 45	8 39 7 27

Maternal Animal	Fetuses Studied Age	Fetuses Studied Length in Mm	Weight in Grams
	Mid Group		
I	39 19	55 5	12.9
J	41 5	55 54 1	16 45 16 25
K	44 6	55 65 55 1	18 65 17 85
L	45 20	60 7 68 9	28.57 27 95
M	46 14	66 3 74 7 70 3	22 77 34 22 30 5
N	48	78 6	40 07
O	49	82 5 68 3	48 35 46 5
	Old Group		
P	50 19	83 15	60 7
Q	53	81 45 75	50 57 44 4
R	55 18	97 1	83 45
S	57	91 9	74 09
T	62 2	102 6 103 6 103 3	84 75 94 75 112

TABLE 2

Responses to	5° C #	5° C %	17° C #	17° C %	27° C #	27° C %	37 5° C #	37 5° C %	47° C #	47° C %	67° C #	67° C %	85° C #	85° C %
Young Group	95	27.46	80	23 12	37	10 69	32	5 20	27	7 80	28	8 09	61	17 63
Mid Group	150	20 56	130	17 81	72	9 86	44	6 03	72	9 86	107	14 66	155	21 23
Old Group	163	23 62	156	22 61	120	17.39	32	4 64	42	6 09	70	10 14	107	15 51

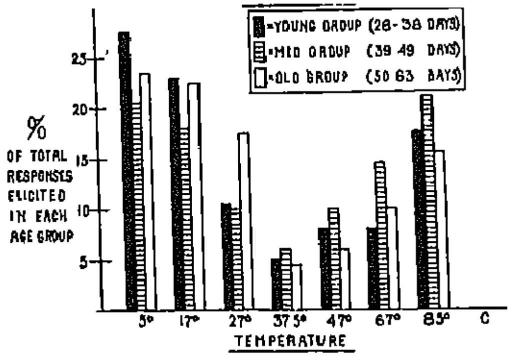
TABLE 3

Group	Possible Responses for any Temperature	5° C. #	5° C. %	17° C #	17° C %	27° C #	27° C %	37 5° C #	37 5° C %	47° C #	47° C %	67° C. #	67° C. %	85° C #	85° C %	Total Possible Responses	Total Responses	Percent
Young	96	52	54 17	44	45 83	28	29 17	17	17 71	19	19.79	16	16 67	36	37.5	672	212	31 54
Mid	78	69	88 46	62	79.49	45	57 69	37	47 44	44	56 41	53	67 95	74	94 87	546	384	70 33
Old	48	41	85 42	40	83 33	35	72 92	20	41 67	15	31 25	34	70 83	37	77 08	336	222	66 07

measurements of the fetuses actually used in this study The adult maternal animals are indicated by letters The age in days and hours since insemination is given next Following this is the crown-rump length and next the weight in grams Inactive fetuses of otherwise active litters are not shown in the table, although they were in each case measured. These measurements were made by G F J. Lehner using a technique previously described by Bridgman and Carmichael (3).

As noted in the table, fetuses from maternal animals *A* to *H* inclusive are here arbitrarily taken to constitute what is called a "young" group, fetuses *I* to *O* inclusive a "mid" group, and *P* to *T* inclusive an "old" group

Table 2 presents the absolute number of responses in each of the three age groups defined in Table 1 for each temperature used. In this table the total of the responses released by stimulating any of the six spots on the body is given Thus, if a stimulus to the ear releases a twitch of the pinna and also leads to palpebral closing,



RESPONSES OF GUINEA-PIG FETUSES TO THERMAL STIMULI

FIGURE 1

two responses are recorded. This table further presents the percentage of the total number of responses actually released at each of the three age groups by each of the temperatures used. A bar graph showing these percentages is presented as Figure 1. This table and graph thus show the relative effectiveness of the specific temperatures used as stimuli *at each of the three periods studied*. It does not directly show the different sensitivities of the receptors at the levels of the three age groups.

Table 3 shows a summary of the total number of responses possible as released by once stimulating each receptor area studied at each temperature used. This table also gives the responses actually elicited by stimulating these areas with the special stimuli as noted. In this table, no matter how complex a response might be, it is merely recorded as one. Thus, to use the same illustration given above, and to illustrate the difference in recording for Tables 1 and 2, if stimulation of the ear releases a twitch of the pinna and palpebral closing, one response only is recorded, that is, the cutaneous area of the ear is here capable of eliciting some response or no response, and only that fact is recorded in Table 3. The total number of *possible* responses can be determined using this method of counting at each of the three periods and at each temperature. This figure is given in the table as is also the percentage of such responses actually elicited at each of the age groups. In general, this table may be taken to show one index of the change in temperature sensitivity which occurs with increasing gestation age.

Table 4 gives one aspect of the character of the responses elicited

TABLE 4

Group	Specific Responses No	Specific Responses % of Total	General Responses No	General Responses % of Total	Total Responses
Young	153	54.64	127	45.36	280
Mid	345	62.96	203	37.04	548
Old	209	58.22	150	41.78	359

by all temperatures at each of the age groups. From the protocols the authors decided whether, in their opinion, each recorded response was to be classified as specific or general. For this purpose, the term "specific" was used to characterize a reflex-like response which is limited in area to what may be roughly called a single response system and ordinarily involving the structure stimulated or some near-by structure The term "general," on the other hand, is used to characterize a response which involves more than a single response system, as defined above. Thus, drops of liquid on the vibrissae which are followed by a twitch of the muscles of the vibrissae only is called specific, but similar stimulation followed not only by a twitch of the vibrissae but also by a twitch of the ear musculature would be defined as general The total number of general and specific responses recorded at each age group and the percentage at each age group are given in this table

Table 5 shows the relative sensitivity of the vibrissae area, the ear, the shoulder, the rump, the fore-paw, and the hind-paw throughout the total fetal period studied as demonstrated by all temperature-released stimuli This table was made by considering responses in the same way in which they were defined in constructing Table 3. It will be noted that the data of this table do not refute the so-called law of cephalo-caudal development or that of proximal-distal development, but, on the other hand, the results do not give any striking confirmation of these generalizations

IV. CONCLUSIONS

The facts presented in the tables and discussion just given, as well as other facts not so fully presented, but available in the written protocols of the experiment which cannot be given here, make possible the following tentative statements concerning the development of temperature sensitivity in the fetal guinea-pig. These statements must be taken as tentative because the series of fetuses used, although representative, was not large enough to justify complete assurance in regard to the generality of the phenomena described.

(1) Temperature stimuli, as defined in this paper, are effective in releasing responses during most of the motile fetal period of the guinea-pig.

(2) At each of three fetal development periods, as these periods were established for the purposes of this study—young, mid, and old—the greater the difference between the temperature of the stimulus and the physiological zero of the organism, the greater the relative number of responses released by that stimulus.

(3). At the youngest ages studied, there appears to be a slight tendency for cold stimuli to be relatively more effective than warm

TABLE 5

Total Possible Responses	Vibrissae #	Vibrissae %	Ear #	Ear %	Shoulder #	Shoulder %	Rump #	Rump %	Fore Paw #	Fore Paw %	Hind Paw #	Hind Paw %
259	180	69.50	155	59.85	147	56.76	106	40.93	133	51.35	97	37.45
Possible Responses—Young Group—112	51	45.54	39	34.82	40	35.71	33	29.46	37	33.04	12	10.71
Possible Responses—Mid Group—91	83	91.21	71	78.02	72	79.12	46	50.55	61	67.03	51	56.04
Possible Responses—Old Group—56	46	82.14	45	80.36	35	62.5	27	48.21	35	62.5	34	60.71

stimuli if certain assumptions of equality are made concerning units of measurement.

(4). During the "young" period of fetal life, as defined above, there appears to be development of temperature sensitivity Sensitivity is greater, that is, in the mid period studied than in the initial period There is no corresponding increase during the other periods studied. In the last period, the growth of the insulating hair coat unquestionably modifies the effectiveness of the stimuli as applied in this investigation

(5). There is no great change in the percentage of specific and general responses released by the stimuli used at any of the three periods studied, nor does the effective intensity of stimulation (as measured by the greater degree of difference from physiological zero) change in a definite manner as development progresses in its efficacy in releasing either general or specific responses There is a slight preponderance of specific over general responses, as these two words are defined in this paper in each period, but this finding can hardly be considered as statistically significant.

(6) There is some shift in the relative sensitivity of the six areas stimulated during fetal development These data may be considered as supporting the generalization made by previous investigators that the development of sensitivity spreads from the cephalic to the caudal regions of the body and from the proximal to the distal regions of limbs

REFERENCES

1 Avery, G T Responses of foetal guinea-pigs prematurely delivered *Genet Psychol. Monog.*, 1928, **3**, 245-331

2 Blanton, M. G The behavior of the human infant during the first thirty days of life. *Psychol Rev*, 1917, **24**, 456-483

3 Bridgman, C. S, & Carmichael, L An experimental study of the onset of behavior in the fetal guinea-pig *J Genet Psychol*, 1935, **47**, 247-267.

4 Canestrini, S Über das Sinnesleben des Neugeborenen *Monog a d Gesamtgeb. d. Neur u. Psychiat*, No 5, Berlin: Springer, 1913 Pp 104.

5 Carmichael, L. Origin and prenatal growth of behavior In *A handbook of child psychology*, 2nd ed, rev, C Murchison, Ed. Worcester, Mass: Clark Univ Press, 1933 Pp 31-159

6. ———. An experimental study in the prenatal guinea-pig of the origin and development of reflexes and patterns of behavior in

relation to the stimulation of specific receptor areas during the period of active fetal life *Genet Psychol Monog*, 1934, **16**, 338-491

7 GENZMER, A Untersuchungen uber die Sinneswahrnehmungen des neugeborenen Menschen. (Dissertation, 1873) Halle Niemeyer, 1882 Pp 28

8 JENSEN, K Differential reactions to taste and temperature stimuli in newborn infants. *Genet Psychol. Monog*, 1932, **12**, 361-479.

9 KRONER, T Über die Sinnesempfindungen der Neugebornen Breslau Barth, 1881 Pp 14

10 KUSSMAUL, A Untersuchungen uber das Seelenleben des neugeborenen Menschen Leipzig Winter, 1859 Pp. 40

11 LOCKE, J Essay concerning human understanding (1690) Philadelphia Kay & Troutman, 1849 Pp 523

12 PEIPER, A. Die Hirntatigkeit des Sauglings Berlin Springer, 1928 Pp 102

13 PRATT, K C The neonate In *A handbook of child psychology*, 2nd ed rev, ed by C. Murchison. Worcester, Mass Clark Univ Press; London Oxford Univ Press, 1933 Pp. 163-208

14 PRATT, K C, NELSON, A K, & SUN, K. H The behavior of the newborn infant *Ohio State Univ Stud, Contrib Psychol*, 1930, No 10 Pp ix + 237

15 PREYER, W Specielle Physiologie des Embryo Untersuchungen uber die Lebenserscheinungen vor der Geburt Leipzig Grieben, 1885 Pp 644

University of Rochester
Rochester, New York

Brown University
Providence, Rhode Island

SERIAL LEARNING II THE BI-DIRECTIONAL GOAL GRADIENT IN THE ENDLESS MAZE*

University of Colorado

KARL F MUENZINGER, CLAUDE C DOVE, AND
ARTHUR H BERNSTONE

Employing a new type of maze, the experiment here reported presents evidence to show that the goal gradient is bi-directional

The study was suggested by Thorndike's fundamental discovery that "the strengthening influence of a reward spreads to influence positively not only the connection which it directly follows and to which it may be said to belong, but also any connections which are near enough to it" (14, p. 3) It occurred to us that the goal gradient might be a special case of the "Thorndike effect," as we have called this phenomenon (10), that the form in which it is usually observed, namely as a gradient of errors increasing from the goal backwards, might be only one-half of the influence of the goal, and that if only a suitable method were used an additional forward effect might be demonstrated

We therefore devised an "endless maze" which permits the animal to go from goal to goal without interruption, thus making it possible for the goal response to "spread its effect," to use the Thorndikian terminology, both forwards and backwards

In the orthodox maze an animal is placed in the starting compartment and then permitted to find the goal. After having reached the goal it is moved back to the starting point by the experimenter or some mechanical device This method interposes a spatial as well as a temporal break between the achievement of the goal and the resumption of the maze learning activity, here the animal runs from *start to goal,* where "start" means a radical change in activity In the situation to be presented the animal runs through the maze as usual But after the goal has been reached and the food eaten, the running through the maze is resumed without interruption, here the animal may be said to run from *goal to goal*

---

*Received in the Editorial Office on October 16, 1936

there had never been any food. This constituted the starting compartment. The door leading out from the next food box was blocked. A rat, after having been placed into the starting box, ran through the maze and had to be lifted out of the goal box. He was then returned to the home cage and given another trial the next day. Thus a rat in this group was given one trial per day for 24 days.

In order to eliminate any possible effect of the particular pattern of correct or unlocked doors used in a maze, six different patterns were employed as follows:

1 — 3 — 1 — 4 — 2
2 — 1 — 3 — 2 — 4
3 — 4 — 2 — 3 — 1
4 — 2 — 4 — 1 — 3
4 — 2 — 3 — 1 — 2
2 — 4 — 1 — 3 — 1

By a pattern designated like the first one we mean that the correct door in the first choice compartment was the first from the left in the direction of travel, and in the following compartments the third, the first, the fourth and the second respectively. The same number of animals in each group learned a particular pattern. Each of the first five patterns was learned by five animals, and the last pattern by two animals in each group.

The 54 rats, of Wistar strain, were about 90 days old at the beginning of the experiment. They were fed for one half hour on a balanced diet after the completion of the day's task.

Results

A. *Group Tendencies.* The errors made at each one of the five choice points are as follows: For the goal-to-goal group:

on choice points	I	II	III	IV	V
total errors	543	668	790	773	534
average per rat.	20.1	24.7	29.3	28.6	19.8

For the start-to-goal group:

total errors	759	716	690	701	466
average per rat:	28.1	26.5	25.5	26.0	17.3

In this compilation several errors at one door in any one trial are counted as one error In the record of the goal-to-goal group the errors made in the four identical mazes of the apparatus are added together, that is, they are treated as if they had all been made in one maze. The results are also presented graphically in Figure 2

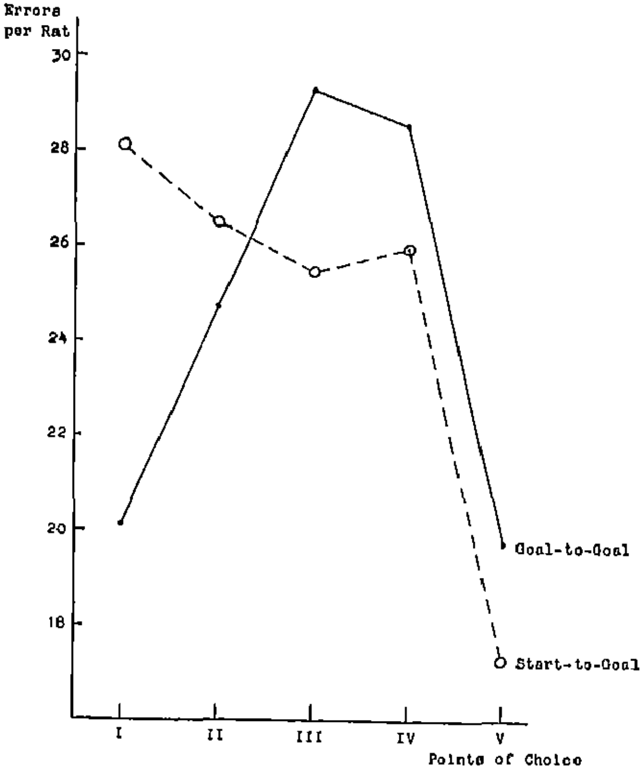


FIGURE 2

ERROR GRADIENTS

The solid dots represent the record made by the goal-to-goal group, the open circles that of the start-to-goal group

The difference between the two groups is striking The control, or start-to-goal, group presents the usual picture in maze learning by animals The greatest number of errors is made near the starting point and the smallest near the goal. This is the orthodox backward elimination of errors The break in the gradient between points III and IV may be due to the patterns used or to other factors; in any case it does not alter the general gradient. (Note also the high number of errors on point IV in the other group One is reminded here of the forward displacement of the point of maximum repetitions in the learning of rote series by human subjects)

On the other hand, the goal-to-goal group shows a forward-backward elimination of errors It has a low number of errors not only at point V just preceding the goal, as one would expect, but also on point I, *the point immediately following the goal* One might speak here of two gradients, one preceding and one following the goal We prefer to consider these two error gradients as belonging together as the manifestation of a single "bi-directional" goal gradient In this phrase the term "bi-directional" is to indicate that the slopes of the gradient before and after the point of reference have opposite signs.[1]

In comparing the two groups one should also note that their records in the last part of the maze run parallel, and that the record of the start-to-goal group is lower than that of the other group as one would expect in the case of distributed as compared with massed practice However, in the first part of the maze the start-to-goal group, despite its distributed practice, starts with a much higher number of errors than the goal-to-goal group The respective averages are 28 1 and 20 1 errors per rat, the difference between these averages is 4 1 times its standard deviation

B *Consistency within the groups* The individual differences between the rats in each group may be expressed in terms of the percentages of the individuals that conform to the general tendencies revealed by the total errors or averages. The average deviations of the averages for each choice point would not be significant in revealing the consistency we desire to know because they would not indicate the error gradients of the individual rats The following summary contains the percentages of rats that show an increase or decrease of errors from point I to the average of points II and III, and from the average of points III and IV to point V. That is, a rat's errors on point I was compared with the average of its errors on the next two points, and its average of points III and IV with its errors on point V

---

[1] Note that the curve in Figure 2, representing the goal-to-goal group, is in striking agreement with a curve of error gradients which Hunter suggested on the basis of the phenomena of backward and forward conditioning (6, p 306, Figure 3) It may be said that Hunter's theoretical analysis predicted our experimental results

From point I to $\frac{\text{II} - \text{III}}{2}$, and from $\frac{\text{III} - \text{IV}}{2}$ to V.

		From point I to $\frac{\text{II} - \text{III}}{2}$	from $\frac{\text{III} - \text{IV}}{2}$ to V
goal-to-goal group	increase	*81.5%* (22)	11.1% (3)
	decrease	18.5% (5)	*81.5%* (22)
	same	0	7.4% (2)
start-to-goal group	increase	33.3% (9)	14.8% (4)
	decrease	*59.3%* (16)	*85.2%* (23)
	same	7.4% (2)	0

The numbers of animals are given in parentheses The percentages in accord with the group tendencies are in italics. In the goal-to-goal group there is a high consistency as between individuals In 81.5 per cent of the rats there is an increase of errors at the beginning of the maze, and in a similar percentage there is a decrease towards the end of the maze The animals of the start-to-goal group are just as consistent towards the end of the maze, but at the beginning they are less so; 85.2 per cent of this group have a decrease of errors at the end, and only 59.3 per cent at the beginning This reduction in consistency might perhaps be expected with an increased distance from the goal

To summarize the results: The error curve of the goal-to-goal group shows a forward-backward elimination of errors, while that of the start-to-goal group shows a backward elimination only Individual differences were small with one exception, immediately after the start in the latter group

DISCUSSION

It is not our intention to review here the literature on the goal gradient We shall merely discuss a possible explanation of the forward-backward elimination of errors in the goal-to-goal situation

Before doing so we should like to point out that our results support the claim of many writers that in the orthodox method of maze learning by animals there is a tendency to eliminate errors in a backward order from the goal We also agree with Washburn

(16, p. 311), Spence (13), Buel (1) and others that many other factors are present in maze running which may mask or obscure this backward learning In our own maze a disturbing factor seems to be the existence of four doors and five choice points necessitating the repetition of a door in the pattern of correct doors

A *Hull's goal gradient hypothesis* The question we must ask here is: Is Hull's explanation of the uni-directional goal gradient (backward learning) applicable to the bi-directional goal gradient (forward-backward learning) in maze learning by animals? Hull's hypothesis (4) has a twofold basis. First, it is assumed that the excitatory tendencies for maze responses with the goal drive as their persisting stimulus decrease with their distance from the goal so that at a point of choice the excitatory tendency to take the right turn is stronger than the one to enter the blind alley. Second, there is experimental evidence to show (17) that the discrimination between two distances to the goal is subject to Weber's law From these two premises it follows that at a choice point nearer the goal the discrimination between two excitatory tendencies (to take the right turn or enter a blind alley) is made more readily than at a point farther away, thus resulting in the backward elimination of errors.

This ingenious explanation fits the facts of the start-to-goal situation where we have a drive to the goal and a backward elimination of errors It does not fit the facts of the goal-to-goal situation where we get a forward-backward elimination of errors although we must assume the same goal drive to exist It is because Hull's hypothesis rests on a goal-directed drive that it is inadequate to account for the bi-directional goal gradient If in the goal-to-goal situation a goal-directed drive is present as the persisting stimulus for maze responses and if the animal discriminates distances according to Weber's law then we should get a backward elimination of errors even in the first part of the maze, a deduction contrary to our experimental results. This raises the further question of the validity of this hypothesis even as regards the uni-directional goal gradient

Hull himself, it should be pointed out, suggests the possibility of a bi-directional gradient (4, p. 39), but he does not indicate how it may be generated, unless he intends his entirely distinct hypothesis regarding the learning of rote series (5) to serve this purpose We would then have one hypothesis accounting for the backward elimination of errors and another one for the forward-backward elimination, which would be an unsatisfactory state of affairs.

Miss Washburn who was the first to suggest an explanation of the backward elimination of errors by means of a "gradient of excitation" based on a goal-directed drive, explicitly states that "it is the putting of an end to the drive" at the goal that makes the mechanism possible. "The process goes thus backward rather than forward, setting in readiness the last movements of the maze rather than the movements made by the rat just after it had finished eating" (16, p. 378).

These explicit statements definitely rule out the possibility of a forward elimination of errors in the first part of the maze in the goal-to-goal situation. For whenever we have a drive in one direction, a goal, and the discrimination of distances to the goal according to Weber's law, the elimination of errors must proceed in the opposite direction.

It would seem then that the goal gradient hypothesis as developed by these writers is inadequate to account for the forward-backward elimination of errors, unless it is extended so as to assume a gradient of excitation that grows weaker in a direction forward from the goal as well as backward. That such a gradient exists is indicated by the Thorndike effect.

B. *The goal gradient as a special case of the Thorndike effect.* Thorndike has demonstrated experimentally (14) that in serial learning where a successful response is preceded and succeeded by unsuccessful ones, not only is the tendency to repeat the successful response strengthened but also the tendency to repeat the unsuccessful ones; and this strengthening of the contiguous connections diminishes with the distance from the successful one both forward and backward.

It seems to us that the forward-backward learning in the goal-to-goal situation is a special case of the Thorndike effect. There is more here than a superficial resemblance between the error curve shown in Figure 2 and a graph representing the repetitions of wrong responses contiguous to a right one in Thorndike's experiment (14, p. 55). In an analysis made by one of the authors (Dove, 2) it was found that the uniformity with which the same error occurs at the same choice point in two successive trials of the goal-to-goal group is higher near the goals than midway between them. We have here a special proof that the Thorndike effect occurs in maze running by rats just as it does in serial learning by human subjects

In our attempt to account for the bi-directional goal gradient in the endless maze with white rats as subjects we may therefore assume as our first premise that the excitatory tendencies for maze responses decrease forward as well as backward from a goal, where the supremely successful response occurs (Thorndike effect)

As a second premise we assume that any variation in the excitatory tendencies in a choice situation affects such a situation as a unit Our basis for this assumption is an experiment in which shock for the right response accelerated learning a black-white discrimination as much as the usual shock for the wrong response We concluded at that time that such an effect could not be expressed in terms of single responses but only in terms of the choice situation as a whole (9, p 274) If we now regard each choice compartment in the endless maze as a unit, any factor affecting the rate of learning in any part of the maze should affect each choice situation as a unit

As a third premise we assume that an increase in the excitatory tendencies in a choice situation is accompanied by an increase in the sensitivity of the cues to be discriminated Increase in such sensitivity leads to an acceleration of learning (9)

From these three premises it would follow that learning immediately before and after a goal takes place at a greater rate than midway between two goals

This account of a bi-directional goal gradient does not explain the Thorndike effect itself. Whether or not such an explanation can be given in terms of forward and backward conditioning, somewhat along Hunter's important analysis, we hope to discuss at some future occasion

As to the nature of the goal response, especially as regards Thorndike's insistence on its satisfying aspect, it is interesting to speculate whether or not the mere presence of food in the starting compartment of an old type maze would impart goal properties to the start, that is, create a forward elimination of errors in the first part of the maze.

C *The Bi-directional goal gradient in human learning* We cannot conclude without emphasizing the discrepancy between human and animal learning In human learning of rote series (3, p. 624, 11) and mazes (15) forward-backward elimination of errors is the rule. Experiments conducted in our laboratory (8) have shown that massed and distributed practice in human maze learning does not produce results comparable to the difference between the goal-to-goal and start-to-goal situations in animal learning as we had expected; both methods of practice with human subjects produced forward-backward elimination of errors. Further analysis showed that the chief reason for such an error curve was symbolic behavior. Responses at the start and end of the maze were verbally fixed at the beginning of training; other responses were added during succeeding trials so that finally the center was approached from both ends. Here then we have a factor in human learning which simulates the manifestation of the Thorndike effect in the goal-to-goal situation in animal learning.[2] On the other hand, the Thorndike effect can be demonstrated very readily in human learning as occurring independently of verbalization. We therefore suggest that the Thorndike effect is one of the fundamental principles of animal and human learning and that masking factors may in either case obscure its manifestations.

SUMMARY

Two groups of rats, hunger-food motivated, were given the same number of maze trials but under different conditions.

The experimental or goal-to-goal group was given 24 trials on one day by means of the endless maze. This apparatus consisted of four identical mazes arranged as the sides of a square with the food boxes at the four corners. As an animal left a food box it went directly into the next maze. Each food box was baited again as soon as an animal had left it. By going around the square six times a rat got the equivalence of 24 trials in the same maze pattern. The record made by this group showed a forward-backward elimination of errors, that is, the first and last correct choices were learned before the middle ones.

The control or start-to-goal group was given one trial per day for 24 days in one of the mazes. The record made by this group showed the typical backward elimination of errors.

---

[2] Still other factors may be present in the case of incidental learning reported by Lepley (7). Here a daily roll call in a large college class produced a gradient of decreasing correctness of names of classmates both forward and backward from a particular student's own name.

It is maintained that the forward-backward elimination of errors in the goal-to-goal situation is a manifestation of the complete goal gradient, which is bi-directional in its fundamental form The backward elimination of errors in the start-to-goal situation is a manifestation of only the first half of the gradient, the other half is absent because the experimental method interposes a radical break in activity between the goal response and the next maze response

Hull's goal gradient hypothesis, which was formulated with a uni-directional gradient in view, does not seem adequate to account for the bi-directional goal gradient This form of the gradient is considered to be a special case of the Thorndike effect.

A hypothesis is roughly outlined to account for the bi-directional goal gradient

Conflicting experimental material from human learning is explained as the effect of masking factors

REFERENCES

1 BUEL, J Differential errors in animal mazes *Psychol Bull*, 1935, **32**, 67-99

2 DOVE, C C The effect of different degrees of success and failure upon uniformity and variability of behavior. *Comp Psychol Monog* (In press.)

3 EBBINGHAUS, H. Grundzuge der Psychologie I 1905

4. HULL, C L The goal gradient hypothesis and maze learning *Psychol Rev*, 1932, **39**, 25-43

5 ——— The conflicting psychologies of learning—a way out *Psychol. Rev*, 1935, **42**, 491-516.

6 HUNTER, W. S Basic phenomena in learning *J Gen Psychol*, 1933, **8**, 299-317

7 LEPLEY, W M A gradient in incidental learning *J Exper Psychol*, 1935, **18**, 195-201

8 MARTIN, D The goal gradient in massed and distributed learning of human subjects *University of Colorado Thesis*

9 MUENZINGER, K F Motivation in learning I Electric shock for correct response in the visual discrimination habit *J Comp Psychol*, 1934, **17**, 267-277

10. MUENZINGER, K F, & DOVE, C C Serial learning I Gradients of uniformity and variability produced by success and failure of single responses *J Gen Psychol*, 1937, **16**, 403-413

11 ROBINSON, E S, & BROWN, M A Effect of serial position on memorization *Amer. J Psychol*, 1926, **37**, 538-552

12 RUCH, F L Goal direction orientation, generalized turning habit, and goal gradient as factors in maze learning in the rat *J Comp Psychol*, 1934, **17**, 225-232

13 Spence, K W The order of eliminating blinds in maze learning by the rat *J Comp Psychol*, 1932, **14**, 9-27

14 Thorndike, E L An experimental study of rewards *Columbia University Contributions to Education*, 1933, No. 580

15. Warden, C J Primacy and recency in maze learning *J Exper Psychol*, 1924, **7**, 98-116

16 Washburn, M F The animal mind New York Macmillan, 1908 Fourth Edition, 1936

17 Yoshioka, J G Weber's law in the discrimination of maze distance by the white rat *Univ Cal Pub Psychol*, 1929, **4**, 155-184

The University of Colorado
Boulder, Colorado

HOW RATS LEARN THE SIMPLE ALTERNATION PROBLEM IN A TEMPORAL WATER MAZE*

Ohio State University

SOL EVANS

STATEMENT OF THE PROBLEM

In the *spatial* maze the white rat readily learns a simple alternation series of almost any length. (A spatial maze is an ordinary bidimensional maze in which the subject is not required to repeat any of the various parts of the maze pathway, and it is therefore a maze in which the subject is required to make only one kind of turn—either right or left—at each choice point.) Furthermore, Hunter (13) has shown that the rat somewhat easily learns a rather long series of simple alternation turns in *a spatial maze which has many of its spatial characteristics largely masked: a tridimensional maze.* Moreover, Carr (1) and Hunter (14) have shown that the rat, with considerable difficulty, learns a simple alternation series (*rl* and *lr*) in *spatial complexes which have many of the characteristics of the temporal maze.* (A temporal maze, Figure 1, is one in which the subject is required to repeat some or all of the parts of the maze pathway in new combinations; it is therefore a maze in which the subject is required to turn in one direction upon arriving at the choice point the first time, and in the opposite direction upon arriving at the choice point the second or some subsequent time.) Finally, Hunter (12) found one rat who mastered a simple alternation series in *a temporal maze.* Six of the seven rats used by Hunter in this test failed to attain mastery. The series used was a long one: *lrlrlr.* The rats were given from 18 to 65 trials each. All of the rats became oriented toward the right.

The success of this one rat on a long series of simple alternation turns in the temporal maze suggested to the writer that perhaps most rats might be able to learn the shortest type of simple alternation series, namely, *rl* or *lr*, and the fact that all of Hunter's rats soon became oriented toward the goal suggested that the simple alter-

---

*Received in the Editorial Office on October 17, 1936.

nation problem might provide a somewhat persistent blind alley response, the evolutionary stages in the elimination of which might become manifest as a function of practice The simple alternation problem was chosen for investigation because it gave promise of providing a somewhat persistent error which the rat, with practice, might finally eliminate

Hunter (12) invented the temporal maze This maze may be used for a great variety of problems: simple alternation (*rl* or *lr*), double alternation (*rrll* or *llrr*), *etc*, and all combinations of these problems or of their various parts Using the temporal maze, Hunter (12, 13, 14) has made a number of excellent studies of the progress made by white rats on the double alternation problem Hunter's main interest, however, seems to have been that of testing for the presence of symbolic processes He was, therefore, interested in finding out whether or not the rat can establish the responses required by the double alternation problem, and not in discovering the evolutionary stages in which a blind alley response becomes a correct response

The problem of this investigation was to discover *how* the white rat learns the simple alternation problem in the temporal maze To do this it is necessary to (*a*) find out how the blind alley response becomes established, (*b*) trace the stages by which this blind alley response is eliminated, and (*c*) discover how the correct response becomes fixed The second aspect of the problem, that of eliminating the blind alley response, deserves special attention Difficulty in eliminating the cul-de-sac prevents an early solution of the problem Throughout practically the whole learning series cul-de-sac elimination is *the* problem of the temporal maze test.

THE MAZE

Hunter's temporal maze was used in a preliminary test, but such great difficulty was experienced in getting the rat to complete a trial in any reasonable length of time in the initial trials that a change of technique became imperative Hunter (12, 13, 14) fed his rats after each completed response, and was thus able to keep them fairly active, and moving in directions somewhat conducive to a solution of the problem The writer, however, desired to avoid feeding during the trial, hoping thus to avoid points of emphasis in the maze so that the performance might be of a more unitary nature with respect to the final goal or exit at the end of each trial. A means was therefore sought whereby the rat could be kept active without resorting to feeding during the trial. The tests conducted by Ruch (18), MacFarlane (16), MacGillivray and Stone (17), Wever (19), Hack (7), and Dunn (3) in which the rat was forced to swim, suggested the use of water in the temporal maze. Accordingly, a temporal water maze was introduced for trial and then used throughout the experiment.

For a number of reasons it now seems most fortunate that this decision was made. First, the use of water kept the rat active, and

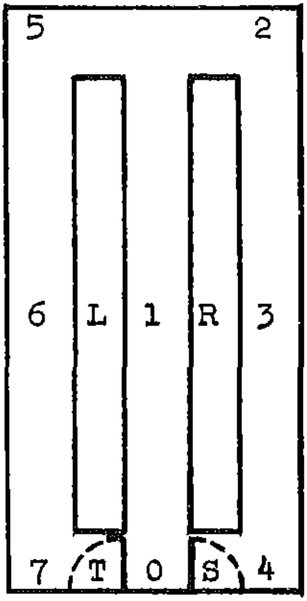


FIGURE 1

A Temporal Water Maze

In the simple alternation problem (*rl*) the animal starts at *O*, and is required to traverse the length of central alley *1*, turn right, follow the direction indicated by the numerals *2, 3, 4*, and pass through door *S* (opened at the appropriate moment) back to *O*, then traverse central alley *1* again, but this time turn left, follow the direction indicated by the numerals *5, 6, 7*, through door *T* (also opened at the appropriate moment) to *O* for the third time, from which position it is allowed to escape by way of a ladder which is lowered just as it arrives.

thus, in most cases, insured the completion of a trial within a reasonable length of time. Secondly, the water seems to have been the factor which brought about the very decisive orientation of the rats toward the goal in the early trials of the experiment, thus providing an opportunity for studying the manner in which persistent errors are eliminated, i.e., how blind alleys are shunted out. And finally, by greatly eliminating inertia, the use of water provided an opportunity for the appearance, and therefore, for the study of many types of hesitation responses, which do not become explicit when the rat is walking on a floor.

The maze (Figure 1) consisted of a heavy galvanized iron tank (24"x48") in which two galvanized iron boxes, *L* and *R*, were placed as indicated. The two doors, *T* and *S*, were hinged to one end of the tank as shown in the figure. The maze pathway indicated by the numerals 0, 1, 2, 3, 4, 5, 6, and 7 is 5½ inches wide. The water in the maze was maintained at a depth of 6 inches, except in the control tests, and at room temperature about 20° *C*. The water was never all removed at one time (except in the control tests). A large pail of water was removed each day and a slightly greater amount added to keep the water clear and to maintain the level in the maze. The tank was supported by a metal frame 36 inches high. The maze was located in a "dark room" in the middle of the floor and was illuminated from the center of the ceiling by a 150-watt light provided with an opaque globe. The maze was separated from the experimenter by a heavy black canvas screen, 6'x8'.

THE RATS

Fourteen white rats 40 days old were used in this experiment. There were 6 males and 8 females. Between trials the rats were kept in cages in a well-lighted laboratory adjoining the room in which the experiment was conducted. The rats were fed on a diet of dog biscuit and various kinds of green material. Fresh water was available for them at all times.

METHOD

The Experimental Procedure. The experimenter started each trial by reaching through a flap in the screen and placing the rat in Section *O*, the maze doors *T* and *S* being in the position indicated in Figure 1. Through a window in the screen he observed the rat traversing the maze, and by means of handles which projected through the screen he operated the doors to comply with the conditions of the problem In the regular trials each rat was given two trials daily, six days per week, one in the morning and one in the afternoon. Escape from water was the incentive used As soon as a trial was finished, the rat was returned in a carrier to his living cage

In the simple alternation problem (*rl*) the rat started at the position *O* (Figure 1) and was required to traverse central alley *1*, turn to the right, follow the direction indicated by the numerals *2, 3,* and *4*, and pass through door *S* (which was opened at the appropriate moment) back to *O*, then traverse the central alley again, but this time turn to the left, follow the direction indicated by *5, 6,* and *7,* through door *T* (also opened at the proper moment) back to *O*, from which position he was allowed to escape from the water by way of a ladder which was lowered just as he arrived

Using the numerals 0, 1, 2, 3, 4, 5, 6, and 7 (which designate the various sections of the maze), the path followed by each rat on each trial was recorded Protocols were also made of the many types of behavior that appeared from time to time hesitation, anticipation, retardation, retracing, etc

The time of each trial was taken by means of a stop watch

The criterion of mastery was three consecutive correct trials

The Use of Terms A number of terms were arbitrarily defined for use in this experiment

A "right turn" (*R*) consists in carrying the head, shoulders, and fore-feet from the choice point, on any choice of a trial, into section *2* (Figure 1)

A "left turn" (*L*) is similarly defined, with section *3* substituted for section *2*

A "correct turn" is a right turn when the choice requires a turn to the right, and a left turn when the choice requires a turn to the left To be a "correct turn" the response in question need not be completed correctly

An "error" consists in carrying the head, shoulders, and fore-feet into a cul-de-sac

A "correct trial" is one in which the rat follows the required pattern without making any error Retracing in the required path does not constitute an error

A "response" on any choice of a trial comprises the behavior sequence intervening between the choice in question and the next choice or the end of the trial.

A "correct response" is a response in which there is no error.

RESULTS

The Establishment of Blind Alley Responses (on Choice 1). Fourteen rats were given two trials daily for 100 days in the temporal water maze shown in Figure 1. On the first two trials most of the rats had not yet become oriented toward the goal but turned right as often as left on both choices of the simple alternation problem. However, after the second trial, most of the animals became decidedly oriented toward the left side of the maze, and so turned left on both choices of the problem for many consecutive trials thereafter. All of the animals became oriented to the left after the 8th trial. This direction tendency was so great that all of the rats, after orientation, made ten or more consecutive trials in which they turned to the left on both choices. Half of the animals continued this left-turning habit uninterruptedly for more than 50 consecutive trials. Table 1 shows the number of trials required by each rat to develop this direction habit (i.e., to entirely discontinue turning to the right) on both choices of the problem. The median number of trials required for complete

TABLE 1

Trials Required for Orientation to the Left on Both Choices of the Simple Alternation Problem

	Trials to Orient Left	
Rat	Total	Correct
1	2	1
2	5	3
3	0	0
4	1	0
6	4	1
7	4	3
8	4	4
9	2	1
10	2	2
11	2	0
12	5	2
13	2	1
14	2	0
15	8	1
Median	2	1

orientation was 2, with a range of from 0 to 8. The great persistence of this left-turning habit was one of the most striking features of the behavior of the animals in this experiment.

Each of the rats, after turning left on Choice 1 and traversing the blind alley to Section 7, would persist in its attacks on the closed door, often for minutes at a time, then it would turn and swim into *Compartment 6, 5, or 1, only to return again immediately to Section 7*. These entries into Section 7, and the many re-entries thereto from the other sections of the blind alley and from Section 1, on the first choice of the simple alternation problem (*rl*), are characteristic of the early trials of the experiment. Re-entering the blind alley on Choice 1 of each trial continued for from 20 to 50 trials.

The Elimination of Blind Alley Responses

1. *Elimination of Re-entries.* Between the 20th and 50th trials the rats discontinued re-entering Compartment 7 on Choice 1, but continued to enter it once. This means that each of the rats turned left and entered Section 7 on Choice 1 in each trial, but did not continue to turn back and re-enter this section from Section 6 or 5 on their way out of the blind alley, also they discontinued returning to Section 1, and from there re-entering Section 7.

It thus occurred that after return trips into Section 7 were discontinued on Choice 1, all rats continued to turn left once on each of

TABLE 2

TRIALS REQUIRED TO MASTER THE SIMPLE ALTERNATION PROBLEM (MAKE THREE CONSECUTIVE CORRECT RUNS) IN A TEMPORAL WATER MAZE

Rat	Before 1st Correct Trial	Trials Before Mastery	Correct Before Mastery
1	47	159	15
2	62	135	17
3	66	143	7
4	20	36	6
6	91	160	4
7	30	75	6
8	100	173	14
9	93		
10	12	168	15
11	44	127	5
12	39	39	2
13	58	78	8
14	47	93	9
15	75		
Median	52.5	131	8

the two choices of each simple alternation trial So the path followed by each animal became 0, 1, 5, 6, 7, 6, 5, 2, 3, 4, 0, 1, 5, 6, 7 (Figure 1) This type of response series continued for many trials with all rats. Table 2 (column 1) gives the number of trials required by each rat before he succeeded in inhibiting the left turn on Choice 1 for the first time, and thus in effecting his first correct run. It will be seen that half of the rats who learned the problem required more than 50 trials to achieve a single correct trial The range was from 12 to 100 trials For all rats except 12, this type of response series continued in great predominance until mastery was achieved Rat 12 achieved mastery at the time of attaining his first correct trial but fell back into the habit of turning left on Choice 1 immediately thereafter and continued in this habit to the end of his trial series. Rats 9 and 15 failed to master the problem in the 200 trials and this type of series continued

2. *The Elimination of a Persistent Blind-Alley Response* Since the left turn on Choice 1 was incorrect, and the left turn on Choice 2 was correct, this left-turning habit gave each of the rats a low score on Choice 1 and a very high score on Choice 2 on the trials before mastery was attained. The left main column of Table 3 will help

TABLE 3

PER CENT OF RESPONSES BEFORE AND AFTER MASTERY ON CHOICES 1 AND 2 OF THE SIMPLE ALTERNATION PROBLEM IN A TEMPORAL WATER MAZE

	Before Mastery			After Mastery		
		Choices			Choices	
Rat	Trials	1 %	2 %	Trials	1 %	2 %
1	159	10.1	99.4	41	73.2	100.0
2	135	12.6	99.3	65	86.2	100.0
3	143	4.9	99.3	57	82.5	100.0
4	36	16.6	94.4	164	78.7	100.0
6	160	2.5	99.4	40	62.5	100.0
7	75	8.0	100.0	125	34.4	100.0
8	173	8.1	100.0	27	81.5	100.0
9	200	3.5	99.5			
10	168	8.9	99.4	32	75.0	100.0
11	127	5.5	98.4	73	11.0	100.0
12	39	5.1	97.4	61	4.9	100.0
13	78	11.6	96.2	122	32.8	100.0
14	93	11.8	96.8	107	73.8	100.0
15	200	6.5	98.0			
Median	131	8.5	99.3	63	73.5	100.0

the reader to comprehend the great extent to which each of the rats maintained his orientation to the left in the trials before mastery was attained This table shows the per cent of correct turns, on the trials before and after mastery, made by each rat on each of the two choices of the simple alternation problem Of the rats who mastered the problem, the median per cent correct on Choice 1 was 8 5, whereas the median per cent correct on Choice 2 was 99 3 How very slowly each of the animals learned the correct responses on Choice 1 is shown in Figure 2 (lower series) This graph shows the per cent of correct responses on choice 1 made by each rat in each of 10 consecutive groups of 20 trials (In this graph all curves follow the 0% line or the plain solid line at all points not designated No curves follow the 0% line between the first and the second group)

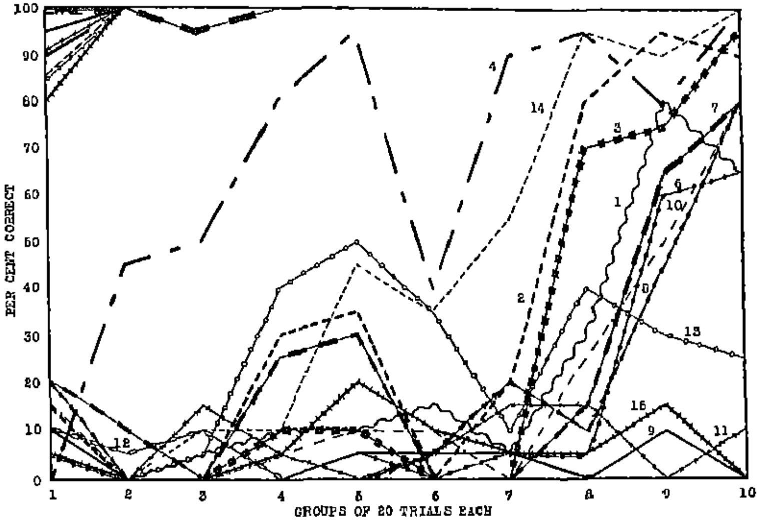


FIGURE 2

Per Cent of Correct Responses Made by Each of 14 Rats on Choice 1 in Each of 10 Successive Groups of 20 Trials

Now contrast the performance on Choice 1 with that on Choice 2 Figure 2 (upper left) shows how very quickly all of the rats learned to respond correctly on Choice 2 (In this graph all curves follow the 100 per cent line or the plain solid line at all points not designated) We have already noted in Table 1 that all of the rats were making consistently correct responses after the 8th trial Figure 2 shows that 3 of the rats responded correctly on Choice 2 in each of the first 20 trials, that 6 rats were correct on 19 out of the first 20; that 2 were correct on 17, and that the remaining one was correct on 16 of the first 20 trials It will thus be seen that of the animals who did make errors on Choice 2 in the first 20 trials, none made less than 80 per cent correct responses As the reader follows through the performance of the rats on the succeeding groups, he will no doubt be struck by the exceedingly great perfection attained and maintained on this second choice by all animals It will be seen that in the 9 succeeding groups of 20 trials each, only one trial by one animal was incorrect This means that only one of the rats made a single error on choice 2 in the next 180 trials. Of 2,470 trials by 14 rats, only one error was made on Choice 2

This great perfection of response attained so quickly by all animals on Choice 2 of the simple alternation problem is a significant fact for the interpretation of observation made on the performance of rats on the double alternation problem and for the explanation of the extension of a learned series of simple alternation responses This point will recur when the problem of explaining the extension of a learned series of alternation arises However, since perfection on Choice 2 was reached almost immediately by all rats, no further analysis of the responses on this choice need be undertaken here An analysis of the changes which occurred on Choice 1 will now be made

In Figure 2 it has been observed that after the beginning of the experiment, all rats soon reached the zero point insofar as correct performance on Choice 1 was concerned. Error on Choice 1 continued to be the predominant type of response with all animals for many trials However, an inspection of these same data will disclose the fact that most of the rats finally did attain mastery of this choice (and hence mastery of the problem) Figure 2 shows that by the end of the 100th trial, rats 2, 4, 7, 13, and 14 had attained a fair degree of perfection on Choice 1 Four of these—4, 7, 13, 14—reached the criterion of mastery before the 100th trial, however, it will be seen that all of these rats fell oft in the accuracy of their performance in the next 20 to 40 trials An explanation for this retrogression in behavior will be attempted later All of these 5 rats attained near-perfection in performance again, except 13 This rat did not again equal its previous record in 200 trials. Errors on Choice 1 were predominant with all rats, except 4 and 14, up to the 160th trial.

Twelve of the rats mastered the problem. Table 2 (middle column) gives the number of trials required by each rat to achieve mastery. For the 12 who mastered the problem, the median number of trials required was 131 and the range from 36 to 173. The last column of the table gives the number of correct trials made by each rat prior to mastery. The median was 8, and the range from 2 to 17 trials.

Rats 9 and 15 did not reach the criterion of mastery within the 200 trials allowed by this experiment.

Table 4 gives the record of each of the rats after mastering the problem. It will be seen from this table that three consecutive correct trials proved to be a good criterion of mastery, as far as serving as an index for the prediction of future success is concerned. Animals 1, 2, 3, 4, 6, 8, 10, and 14 each succeeded on about 75 per cent of all trials after mastery. If conditions were still more rigidly controlled, it seems possible, in this experiment, that the animals might pass from a past record of complete failure to a record of complete success. Column 3 of the table gives the greatest number of consecutively correct

TABLE 4

TRIALS CORRECT AFTER MASTERY ON THE SIMPLE ALTERNATION PROBLEM (*rl*) IN THE TEMPORAL WATER MAZE

Rat	Trials After Mastery Total	Correct	Greatest Number Consecutively Correct
1	41	30	9
2	65	56	22
3	57	47	20
4	164	129	23
6	40	25	8
7	125	43	21
8	27	22	6
9			
10	32	24	9
11	73	8	3
12	61	3	3
13	122	40	3
14	107	79	24
Median	63	35.0	8.5

trials made by each of the rats. With rats 7, 11, 12, and 13 the criterion did not prove to be such a good index to future success. After making 7 consecutive correct runs, rat 7 fell back into a less efficient method of responding, namely, turning left at the choice point, going as far as Section 5, then turning and completing the remainder of the run successfully. This type of series predominated for about 80 trials, i.e., until the rat again relearned the correct turn on the 157th trial. From then on his performance was nearly perfect.

Rat 11, after barely reaching the criterion of mastery on the 127th trial, fell back into a type of response similar to the reversion which occurred with rat 7, or the still more primitive type of response of going to Section 7 before completing the trial. These predominant types of response were interspersed with occasional correct runs. This rat did not again reach the criterion of mastery within the 200 trials. Rat 12, after barely reaching the criterion of mastery, on trial 39, fell back into two much more primitive types of response, going either to 6 or 7 on Choice 1, the latter response predominating very greatly. One or the other of these two types of response persisted until the 85th trial, after which Choice 1 carried the animal to 5 on all trials. This animal developed a wry neck after the 100th trial, and was thereafter discarded from the group. This rat did not make another perfect run after the criterion of mastery was reached. Rat 13, after barely reaching the criterion of mastery, rather persistently alternated one or two correct trials with from 1 to 16 trials that were incorrect. There were 16 consecutive trials in which one correct trial was alternated with one incorrect trial. The incorrect responses on Choice 1 always carried the rat into Section 7. However, this rat again barely reached the criterion of mastery on the 191st trial.

A general picture of the frequency of occurrence of the *different possible combinations of right and left turns* may be had by a glance at Table 5. This table gives the per cent of the trials, before and after mastery, which fell into the six possible series of turns occurring on the two choices of the simple alternation problem. The table also stresses the point that has just been made, namely, that before mastery the rats were decidedly oriented toward the left on both choices of the problem.

3. *Cul-Shortening.* It has already been observed that Choice 1 was the only point of difficulty for the rats in the simple alternation problem. This choice was in fact *the* problem for all of the rats

TABLE 5

PER CENT OF TRIALS, BEFORE AND AFTER MASTERY, THAT FALL INTO THE SIX POSSIBLE SERIES OF TURNS OF THE SIMPLE ALTERNATION PROBLEM IN THE TEMPORAL MAZE

	Before Mastery							After Mastery						
		Six Series of Responses							Six Series of Responses					
Rat	Trials	RR %	RL %	LR %	LL %	R– %	L– %	Trials	RR %	RL %	LR %	LL %	R– %	L– %
1	159	6	11.3		88.1			41		75.6		24.4		
2	135	8	13.3		85.9			65		89.2		10.8		
3	143	1.4	6.3	1.4	90.2		7	57		82.5		17.5		
4	36	2.8	22.2	2.8	69.4	2.8		164		80.5		19.5		
6	160		4.4		95.0		6	40		65.0		35.0		
7	75	1.3	8.0	1.3	89.4			125		36.0		64.0		
8	173	6	10.4		89.0			27		81.5		18.5		
9		5	6.5	1.5	91.5									
10	168		29.2	6	70.2			32		75.0		25.0		
11	127	1.6	10.2		88.2			73		12.3		87.7		
12	39		7.7	7.7	82.0	2.6		61		4.9	1.6	93.5		
13	78	1.3	14.1		82.1		2.5	122		33.6		66.4		
14	93	2.2	11.8		84.9		1.1	107		73.8		26.2		
15			8.5	1.5	89.0		1.0							
Median		7	10.3	3	88.2	0.0	0.0			74.4	0.0	25.6		

on this test. All of the rats, for many consecutive trials after the experiment began, made the error of turning left and proceeding to Section 7, the end of the blind alley. It therefore became a point of much interest to discover how this persistent error would be eliminated In keeping the protocols of the trials a record was made of the complete path traversed by each of the rats on each trial It was therefore possible to discover from the records whether the blind alley on Choice 1 was eliminated in an all-or-none fashion or by gradually shortening the trips into the alley. An analysis of the data on this point seemed to be a worthy enterprise, since it appears that no analysis has yet been made of the elimination by rats of blind alleys in the temporal maze

The trials made by each of the rats may readily be divided into 7 different types on the basis of the type of response which immediately succeeded Choice 1, and in conformity with the numbered sections of the maze (Figure 1) The most common type of response on Choice 1 in the trials before mastery was that of making a left turn and going to Section 7, the end of the blind alley, and from that point completing the remainder of the trial correctly. This way of responding was called Type 7 The other types that are analogous to this, except that the rat turns around after going only as far as 6 or 5, are called Types 6 and 5 respectively Three other types are those in which a right turn carries the rat only as far as Section 2, 3, or 4 respectively, from either of which sections the animal returns to the blind alley before finishing the trial correctly. These were called Types 2, 3, and 4 respectively The remaining type is the one we have termed a "correct" trial, and consists of completing the series of responses required by the problem without entering the blind alley at all

Thus a hierarchy of types may be arranged—7, 6, 5, 2, 3, 4, C—each successively superior to the preceding type from the point of view of nearness to the solution of the problem.

Table 6 gives the per cent of the total responses, before and after mastery, that fall into each of these seven types of response made on Choice 1. It will be observed that with all rats most of the trials prior to mastery involved blind alley responses to Section 7 on Choice 1 This was not only the most common type of error but also the most common type of response on this choice in the trials before mastery The median per cent of Type 7 response before mastery was

TABLE 6

Per Cent of Trials, Before and After Mastery, that Fall into Each of the Seven Types of the Hierarchy of the Simple Alternation Problem (*rl*) in a Temporal Water Maze

			Types of Response						
Rat	Trials		7 %	6 %	5 %	2 %	3 %	4 %	C %
1	Before	159	84.9	1.3	1.9	.6	1.3		10.0
	After	41	24.4					2.4	73.2
2	Before	135	65.2	8.9	11.9	.7	.7		12.6
	After	65	9.3	1.5		1.5	1.5		86.2
3	Before	143	90.2	.7	1.4	2.1		.7	4.9
	After	57	17.5						82.5
4	Before	36	72.2			8.3		2.8	16.7
	After	164	19.5			1.2		.6	78.7
6	Before	160	69.4	6.2	20.0	1.9			2.5
	After	40	10.0		25.0	2.5			62.5
7	Before	75	74.7	9.3	6.7	1.3			8.0
	After	125	7.2	3.2	54.4			.8	34.4
8	Before	173	76.9	3.4	8.7	2.9			8.1
	After	27	3.7		14.8				81.5
9	Before	200	(69.0)		(24.0)	(3.5)			(3.5)
	After								
10	Before	168	65.5		5.4	19.0	.6	.6	8.9
	After	32	21.9		3.1				75.0
11	Before	127	79.5		8.7	5.5	.8		5.5
	After	73	43.8		43.8	1.4			11.0
12	Before	39	87.2		2.6			5.1	5.1
	After	61	65.6	3.3	26.2				4.9
13	Before	78	80.7	1.3	2.6	2.6	1.3		11.5
	After	122	66.4			.8			32.8
14	Before	93	84.9	1.1		2.2			11.8
	After	107	26.2						73.8
15	Before	200	(90.0)		(1.5)	(2.0)			(6.5)
	After								
Median	B	131	78.2	1.4	4.0	2.2	0.0	0.0	8.5
	A	63	20.7	0.0	0.0	0.0	0.0	0.0	73.5

78.2. With rats 1, 2, 3, 4, 10, 12, 13, and 14 this type of error also occurred more frequently than any other in the trials after mastery. The median per cent of Type 7 response for the trials after mastery was 20.7.

Now let us consider the manner in which the rats eliminated this blind alley. It will be noted that rat 4 was the only rat who did not shorten the blind alley response on Choice 1, one or more times before mastery was reached. Rat 14 was almost equally clean-cut with his turns, shortening the response only once by turning in 6, and not at all by turning in 5 The other 10 animals who learned the problem shortened the cul-response on a number of trials, by one or both of these types of cul-shortening responses The median per cent of Type 6 responses before mastery was 1.4. With only one rat, No 7, was this the most frequently-occurring cul-shortening response The median per cent of Type 6 after mastery was 0 0 This type appears to be a pre-mastery phenomenon. With no rat was this the most frequently-occurring cul-shortening response after mastery

The median per cent of Type 5 responses before mastery was 4 0. With all rats except 4, 14, and 10, this was the most frequently-occurring cul-shortening response in the trials prior to mastery The median per cent for trials after mastery was 1 6

The size of the last median given does not convey a correct impression of the great persistence of Type 5 responses after mastery with some of the rats Rats 11 and 12, after attaining mastery, were unable to maintain it, and fell back into the more primitive Type 5 response This was by far the most frequently-occurring type of cul-shortening response with these two rats in the trials after mastery They were continuing this type of response at the end of the experiment Rats 6 and 7 also reverted to more primitive types of response with Type 5 greatly predominating With these two rats Type 5 response was not only the most frequent cul-shortening response but also the most frequent error

Responses 2, 3, and 4 were characteristic pre-mastery phenomena, and disappeared with mastery Type 2 is of no little interest because in it the rat is making a new response to an old situation. The rat presents a picture to the observer of being entirely stranded Entrance into the blind alley is inhibited, but what habit has been built up to take its place? While observing this behavior, the experimenter is tempted to become anthropomorphic and infer that the rat is saying to himself, "There is no use going into the blind alley, but where *is* there any place to go?" The median per cent for Type 2 responses before mastery was 2 2, and after mastery, 0.0 Rat 10 was held up on Type 2 in 19 0 per cent of his trials before mastery

Type C, correct responses, became predominant with all rats except 9, 11, 13, and 15 by the 180th trial The median per cent for this type before mastery was 8 5 and after mastery, 73.5

4 *Hesitation Responses* We have just noted that all of the rats, with one exception, engaged in a number of cul-shortening responses

before developing the habit of consistently eliminating the blind alley on Choice 1 The most significant observation made on this series of tests, however, was the appearance and persistence of hesitation behavior accompanying the response undergoing modification Previous to the appearance of the hesitation response, all rats had been hugging the left central alley—anticipating the left turn into the blind alley

Hesitation behavior began to manifest itself in all rats on Choice 1 just prior to, coincident with, or immediately subsequent to the first correct trial (Table 2) Many types of hesitation responses appeared, but for purposes of exposition, they are all placed under three main headings in Table 7 These are called *hesitation proper*, *central alley return*, and *end wall encounter* responses Under the heading, hesitation proper, is included the response of starting left, then going right, and *vice versa*, turning the head left, then going right and *vice versa*, being held on a balance and vacillating back and forth, slackening the swimming speed greatly at the choice point, hitching quickly to the right or left, swerving back and forth in arriving at the choice point, stopping at the choice point for a short or a long time, barely entering the blind alley and turning right, turning left through an arc of 270° and going or starting to the right, and shifting back and forth from one side of the central alley to the other at the choice point *rl*, *lrl*, *rlrl*, *lrlrl*, *rlrlrl*, *lr*, *rlr*, *lrlr*, *rlrlr*, or *lrlrlr*

All of the rats engaged in many of the different varieties of response listed as "Hesitation Proper" and several of the animals engaged in most of the different varieties The variety which occurred most frequently with most animals was the *rl* or *lr* . kind

The other two main types of hesitation response listed in Table 7 are almost self-explanatory A central alley return response was a response in which the rat, after having traversed the length of the central alley from the starting place to the choice point, turned around at the choice point and retraced all or part of the distance back to the starting place before finishing the trip through the maze An end wall encounter was a response in which the rat, without slowing up or turning to the right or left, swam straight ahead and bumped into the end wall of the maze, just beyond the choice point

Table 7 lists the per cent of the total responses (correct and incorrect) made by each rat on Choice 1, before and after mastery, that were characterized by the three different types of hesitation behavior

TABLE 7

PER CENT OF THE TOTAL RESPONSES (CORRECT AND INCORRECT) OCCURRING ON CHOICE 1 OF THE SIMPLE ALTERNATION TRIALS, BEFORE AND AFTER MASTERY, CHARACTERIZED BY THREE DIFFERENT TYPES OF HESITATION BEHAVIOR

	Types of Hesitation Behavior on Choice 1											
	Hesitation Proper				Central Alley Return				End Wall Encounter			
	Before		After		Before		After		Before		After	
Rat	Inc. %	Cor %	Inc %	Cor %	Inc. %	Cor %	Inc. %	Cor %	Inc. %	Cor %	Inc. %	Cor. %
1	11.3	3.8	17.0	75.6	2.5							
2	9.5	1.5	4.5	7.6	1.5					7		
3	9.8	2.1	5.3	42.1								
4			4.3	6			1.2		2.8		6	
6	5.5	1.8	2.5	25.0								2.5
7	12.1	1.3	4.8	14.4	1.3		3.2					
8	17.4	2.3	3.7	14.8	4.0		3.7		1.2			
9	21.5	2.5			3.0							
10	16.1	6.0	6.2	18.8	1.8	6			3.0			3.1
11	26.8		16.3	2.8	8							
12			4.8	3.2								
13	9.0		3.3	1.6					2.6			
14	3.3	1.1	1.9	53.2					9			1.9
15	19.5	2.0			2.5				6.0			
Median	10.6	1.8	4.7	14.6	1.1	0.0	0.0		4	0.0	0.0	0.0

5. *Hesitation Proper.* A glance at the type entitled "Hesitation Proper" will disclose that a majority of the rats hesitated at the choice point more frequently after mastery than before mastery. Nine of the 12 rats (1, 2, 3, 4, 6, 7, 10, 12, 14) who mastered the problem hesitated on a greater per cent of the choices after mastery than before mastery. The most interesting fact, however, is that after mastery a choice upon which hesitation occurred stood a better chance of terminating in a correct response than the same kind of choice before mastery. Before mastery the median per cent of *incorrect* hesitation responses was about six times the median per cent of *correct* hesitation responses. After mastery was achieved the relationship was reversed: the median per cent of correct hesitation responses was then three times the median per cent of incorrect hesitation responses. Thus the rats were not only hesitating more frequently, but the hesitation movements in which they were engaging were terminating more certainly in correct responses. The rats were gradually learning to engage in a variety of activities without disrupting whatever posture was requisite for inhibiting entrance into the blind alley. It may well be said that the rats at first *went the wrong way;* next, they learned *to hesitate and go the wrong way,* they then learned *to hesitate and go the right way.* At the close of the trial series a few rats had reached the point in their training where they merely *went the right way.* Most of the rats, no doubt, would soon have reached the last stage in their training, however, the training series was discontinued too soon to lower greatly their hesitation record.

At first no record was kept of the time consumed in each hesitation response. The responses were of such short duration as to make their temporal importance appear insignificant for purposes of this investigation. With more trials, however, hesitation time increased. It was soon recognized that the duration of these hesitation responses might prove to be a significant factor in this investigation. Accordingly, a *time record was inaugurated for that variety of hesitation response* consisting of one or more shifts back and forth across the central alley at the choice point. This was the most prevalent variety of hesitation response. The median time consumed in this variety of hesitation was 12.2 seconds. There were 23 hesitation responses in which the subjects shifted back and forth 4 times in front of the choice point, and which terminated in correct trials. Four of these 23 responses were by 2 rats before mastery, and the other 19 were by 8 rats after mastery. The median duration of this variety of hesitation was 21.3 seconds.

Another variety of this type of hesitation response should be mentioned. Two of the rats, 6 and 7, in 7.5 per cent and 7.2 per cent respectively of their trials after mastery, engaged in a variety of hesitation responses consisting of swimming to the choice point in the usual manner, and then turning around left as on a pivot through an arc of 270° before proceeding to the right. This was undoubtedly the remnant of a blind alley response which no longer carried the rat into the blind alley. Rat 14 engaged in a somewhat similar kind of hesitation response in over 50 per cent of the trials after mastery. This response consisted in a sudden hitch to the right after turning the head and neck toward the entrance to the cul-de-sac.

6. *Central Alley Returns.* The second type of hesitation behavior, the central alley return response, was exhibited before mastery by nine rats. With eight of these all such responses before mastery were incorrect or blind alley responses. Before mastery only one rat, No. 10, succeeded in making any successful responses after retracing in the central alley. This rat made two such responses after retracing. The median per cent of incorrect central alley retrace responses before mastery for the rats who learned the problem was 1.1, and the range from 8 to 4.0. Rat 8 made 8 such unsuccessful retrace responses before mastery. The median per cent correct retrace responses was 0.0.

If we turn to the retrace responses made by the rats after mastery we have much the same picture. However, fewer rats participated in such responses. Only 3 rats—4, 7, 8—engaged in retracing responses after mastery. The feature that makes these responses similar to the retracings before mastery is that all of them terminated in unsuccessful responses on Choice 1. It appears that central alley retracing is largely a premastery phenomenon, which may occasionally appear in trials after mastery; but that its appearance even in trials after mastery is likely to result in an unsuccessful response on the choice with which it appears.

The almost certain failure which results subsequent to these central alley retrace responses suggests that the limit as well as the length of the transition stages of the cul-shortening process may be an important factor in the rate of maze learning; and that the rat who shortens the alley by the greatest transition steps need not always be the most rapid learner It has already been shown (Table 6) that cul-shortening does not advance by imperceptible steps, but rather by jumps (small or large) Rat 4 passed directly from Type 7 responses to Type 2, i e., shortened the blind alley by its full length on the first transition step Except for one trial the same was true of rat 14 These two rats were by far the best performers in the group On the other hand, rats 9 and 15, whose first transition step was intermediate in length (from Type 7 to Type 5), were decidedly the poorest performers (failed to learn the problem) The rather high per cent of central alley retrace responses in which these last two rats engaged suggests that the limit of the second step of the cul shortening process fell somewhere in the central alley, and resulted in a series of incorrect responses. The point that is being made is that a large transition step is conducive to rapid learning, provided its limit falls at or near the choice point, but that such a step complicates the problem when its limit falls any considerable distance into the central alley Short transition steps (the type occurring with most of the rats) stand a better chance of providing a limit which falls near the choice point. This may be the explanation as to why all of the rats who employed short transition steps sooner or later learned the problem

Rat 2 presents a case in which the cul-shortening process, after passing through a number of transition steps, had apparently extended back into the starting box For many trials subsequent to mastery, this rat remained in the starting box for about one minute before starting on his trip through the maze While in the starting box this rat appeared very tense and inhibited (presented the same general picture that most of the rats did in their hesitation responses at the choice point just prior to rejecting the blind alley), but at the end of his stay there, he began to swim very freely up the central alley while hugging the right side wall Upon reaching the choice point he turned to the right without hesitation Near the time of their attainment of mastery a number of such responses occurred with four other rats

7 *End Wall Encounters* The third type of hesitation response which occurred on Choice 1 was the end wall encounter response Nine rats engaged in this type of response. Eight exhibited the response before mastery Seven of these made incorrect responses on each trial after bumping into the end of the maze Before mastery only one rat, No 2, succeeded in making a correct response on Choice

1, after having bumped into the wall. The median per cent of incorrect end wall encounters occurring before mastery was 4. The median per cent of correct end wall encounters was 0.0. In the end wall encounter responses occurring after mastery there was a somewhat different trend. Four rats participated in end wall responses after mastery. But, whereas end wall encounter responses before mastery terminated in incorrect responses on Choice 1 with almost all rats, the same type of hesitation responses terminated in correct responses after mastery. It seems evident that three of the rats had learned to maintain the posture necessary to inhibit entrance into the blind alley even after crashing into the end wall. One animal, however, still participated in one end wall response which terminated in an unsuccessful response.

It appears from this observation that the end wall encounter response is predominantly a pre-mastery phenomenon, and that with training the rat may learn to engage in such activity without becoming disinhibited.

Reversion. A striking feature of the performance of the rats on the alternation problem was the phenomenon of reversion. After reaching any one of the more advanced types of performance, each of the rats occasionally, and many of them persistently, continued to fall back into a more primitive and less adequate type of solution. This continued to occur on a considerable scale until near-perfection was reached. This fact has been recognized by many investigators (8). The phenomenon is probably to be attributed to the impossibility of attaining absolute control over the stimulating situation while a response is suffering experimental extinction.

Whatever may be the basis for reversion, it was believed to be an important phenomenon for the explanation of the development of mastery of the alternation problem. An investigation into the problem of whether the alternation habit, which had developed through the various stages of the hierarchy, could be forced back into some of these stages through which it had developed was considered a worthy undertaking.

Reference has already been made to the fact that by the end of the 100th trial, rats 2, 4, 7, 13, and 14 had reached a rather advanced stage of mastery. All of these except number 2 had mastered the problem in terms of our criterion. It will be observed (Figure 2) that in the 5th group of 20 trials, rat 2 was correct on 35 per cent, rat 4 was correct on 95 per cent, rat 7 on 30 per cent, rat 13 on 50 per cent, and rat 14 on 45 per cent of the trials. The remaining 9 rats were making scores not much better than at the beginning of the series, but many of them by this time had passed through the lower hierarchy of types and were just on the brink, so to speak, of perfecting the habit.

The alternation habit had appeared so unstable throughout its development that it was believed that the experimental situation provided a good opportunity for a test of the reversion hypothesis. The habit with most rats had developed through a part or all of the hierarchy of stages, and on the 100th trial, there were rats in all of the different stages. Furthermore, all of the rats, even those who had mastered the problem, seemed still to be very strongly oriented toward the goal at all times except at the time, or immediately preceding the time, that Choice 1 was made. These conditions seemed to provide a splendid opportunity for testing the hypothesis. A habit without a hierarchy of stages does not provide such an opportunity. In order to test for reversion, it is necessary to have something to which to revert, and this something must be known.

It was desired that the stimulating situation be changed in order that attendant changes in the behavior, if any occurred, might be observed, but it was not desired that the change be great enough to throw *all the animals back into the most primitive level of the hierarchy*, i.e., Type 7 (Table 6). The modification introduced was that of changing the water in the maze and raising the level 1½ inches.

As has been previously mentioned, it was not the practice in the experiment to remove all of the water from the maze at one time. Only one large pail of water was removed and a like amount replaced each day. However, for this control test, all of the water was removed from the maze, the maze thoroughly washed, and refilled with fresh water at room temperature, but at a level 1½ inches higher than in the preceding trials of the experiment. The rats were then given all of their remaining trials (100) in water at the new level.

The per cent of trials of each of the rats that fell into each of the 7 types of the hierarchy (Table 6) previously discussed was computed on the basis of the 7 trials given subsequent to the introduction of the new feature into the stimulating situation. The data for the 27 trials previous to the introduction of the new stimulus were also computed for comparison. This comparison is made in Table 8.

TABLE 8

PER CENT OF TRIALS THAT FALL INTO EACH OF THE SEVEN TYPES, BEFORE AND AFTER MODIFYING THE STIMULATING SITUATION

		Hierarchy of Types						
Rat		7 %	6 %	5 %	2 %	3 %	4 %	C %
1	Unmodified	88.9		3.7				7.4
	Modified	71.4		14.3				14.3
2	Unmodified	26.0	7.4	40.7				25.9
	Modified	71.4	14.3	14.3				
3	Unmodified	81.5		7.4			3.7	7.4
	Modified	100.0						
4	Unmodified	14.8					3.7	81.5
	Modified	71.4						28.6
6	Unmodified	51.9	14.8	25.9	3.7			3.7
	Modified	57.1		42.9				
7	Unmodified	14.8	7.4	48.2			3.7	25.9
	Modified	14.3		85.7				
8	Unmodified	92.6		3.7				3.7
	Modified	85.7		14.3				
9	Unmodified	51.9		22.2	18.5			7.4
	Modified	71.4		14.3	14.3			
10	Unmodified	25.9		7.4	48.2	3.7		14.8
	Modified	57.1		14.3	28.6			
11	Unmodified	63.0		22.2	11.1			3.7
	Modified	100.0						
13	Unmodified	48.2			3.7			48.1
	Modified	71.4						28.6
14	Unmodified	51.9			3.7			44.4
	Modified	71.4						28.6
15	Unmodified	63.0		7.4	3.7			25.9
	Modified	100.0						
Median	Unmodified	51.0	0.0	7.4	3.7	0.0	0.0	14.8
	Modified	71.4	0.0	14.8	0.0	0.0	0.0	0.0

It will be observed that reversion took place with all of the rats with the possible exception of rat 1. In general, the rats decreased in the per cent of trials falling into the higher levels of the hierarchy and increased in the per cent of trials falling into the lower levels of the hierarchy. Twelve of the 13 animals decreased in their percentage of correct trials. The median per cent correct before disturbance was 14.8, and after disturbance, 0.0. There was a decrease in percentage of trials of Types 3 and 4 with all rats who had been engaging in either of these types, although the number participating in these types was not great enough to bring the medians above 0 0 per cent. Of the 7 rats engaging in Type 2 trials, all had their percentage of trials of this type diminished coincident with the introduction of the disturbing stimulus The median per cent of the last-named trials was 3 7 before disturbance, and 0 0 after disturbance

On the other hand, the per cent of trials in two of the three lower levels of the hierarchy was greatly increased It is true that of the 10 rats who engaged in Type 5 trials, only half of them had their percentages in this type increased incident with disturbances, but there was an increase of the median from 7 4 per cent to 14.4 percent, and with the 5 rats whose per cent of Type 5 trials decreased, there was a greater compensation toward the lower level of the hierarchy in the per cent of Type 7 trials. The per cent of Type 6 trials was increased with one of the three rats who were participating in this type of trial, and became zero with the other two The per cent of Type 7 trials was increased with ten of the 13 rats The other three—6, 7, 8—had their per cent of this type of trial slightly lowered, but this decrease was more than compensated for by the increase which occurred with these rats on Type 5 trials A comparison of the medians before and after disturbance gives a good conception of the increase that occurred in this type of trial The median per cent before disturbance was 51 9, after disturbance, 71 4 It will be observed from Figure 2 that this disturbance was not overcome with most of the animals until after about 40 or 50 trials

This reversion test indicates that the rat was not performing on the basis of starting-box cues, but in accordance with the stimulating situation as it developed while the rat was traversing the central alley

The Extension of a Learned Series Gellermann (4, 5, 6) and others have noted in monkeys and in children a tendency to alternation Husband (15) has shown that human subjects tend to alternate while the rat develops a direction tendency That the rat soon develops the direction habit under certain conditions in the temporal maze, there can be no doubt This is shown in this experiment by the fact that the rat always turns left on both choices of the simple alternation problem for many consecutive trials, before eliminating the blind alley on Choice 1 However, the fact that the second choice was invariably *a turn to the left in spite of the fact that a left* response on the previous trial had been unsuccessful, suggested the possibility that a simple alternation series, after once under way, is a very natural and elementary sequence of responses for the rat to develop

About two weeks after the learning and control trials were finished and the rats had rested for 10 or 15 days, a test of the ability of the rats to successfully extend the simple alternation habit of two choices into an alternation series of four choices was undertaken Twelve rats were used in the test The number of trials given to each of the various rats ranged from 4 to 8

During the period of rest a number of the rats had fallen off in their ability to succeed on Choice 1 of the simple alternation problem However, whether or not the rat was successful on Choice 1, it was allowed the opportunity of extending the series to *four choices*

Table 9 gives the results of the extension test It will be observed that 8 of the 12 rats extended the series with considerable degree of success Four of the rats who successfully extended the series made a better record on the last two choices than on the two choices on which they had been trained The other four did not do as well on the last two choices as on the first two but their reduction in efficiency was not great When failure occurred in the extension of the series

TABLE 9

PER CENT OF CORRECT RESPONSES ON CHOICES 1, 2, 3, AND 4 IN EXTENDING A TWO-CHOICE SIMPLE ALTERNATION SERIES (*rl*) TO A FOUR-CHOICE SIMPLE ALTERNATION SERIES (*rlrl*)

		Choices			
		1	2	3	4
Rat	Trials	%	%	%	%
1	6	16 7	100 0	66 7	100 0
2	8	100 0	100 0	87 5	100 0
3	4	50 0	100 0	75 0	75 0
4	8	100 0	100 0	75 0	100 0
6	7	28 6	100 0		100 0
7	7	28 6	100 0	85.7	100 0
8	5	100 0	100 0		100 0
9	6	83 3	100 0		100 0
11	7	42 9	100 0		100 0
13	5	60.0	100 0	40 0	100 0
14	7	100 0	100 0	42 9	100 0
15	5		100 0	40 0	100 0
Median	6 5	55 0	100 0	41.5	100 0

on any trial by any one of the eight rats, it was probably brought about by a long delay in the starting box, after the second response had been completed the animals persisted in attempts to climb the wall of the maze at this point in their progress. All 8 rats, on all trials in which they proceeded up the central alley immediately after finishing the second (left) response, were successful in making a right response on Choice 3.

Four of the rats failed completely on the problem of extending the series, turning left on both Choices 3 and 4 The explanation of failure on all trials in the case of these four rats is probably the same as that with the other eight on the trials in which they failed Each of the four rats who completely failed in extending the series stopped for long periods of time to attack the walls of the starting box after completing the second response of the series

It is a surprising fact, not that some of the animals failed on all trials to extend the series, nor that all animals failed on some of the trials, but that any of them succeeded in extending any of the trials

It will be remembered that all of the rats for a great number of consecutive trials in the training series responded by going left on *all choices* Now that the rats should reach a median of 41 5 per cent success in extending the series despite the fact that all of them delayed in the starting box after Choice 2 in some or all of their trials is a most striking fact indeed This is the phenomenon which Hunter (9, p 216, 10, p 531, 11, p 1) has called "alternation after success" A rat who has recently traversed a path around one of the rectangles does not so readily repeat the path, but is more prone to follow a path leading in the opposite direction This tendency to alternate bears on the problem of learning by rats of the simple alternation temporal maze It explains why all of the animals were always correct on Choice 2 almost from the very beginning of the series of trials It will be observed that in the test of extending the series, all animals except one were correct on Choice 4 on all trials upon which they were tested

RÉSUMÉ OF THE LEARNING PROCESS

The observations made upon the rats in the learning series, in the reversion test, and in the test on extending a series of simple alternation responses have provided data, it appears, for a somewhat complete description of the development of the simple alternation habit in the temporal maze Such a description will now be attempted

On the first trial the rat is neutral with respect to the various sections of the maze and therefore with regard to Choice 1 of the problem He will therefore be as likely to turn in the one direction as in the other at the choice point Whether he turns to the right or to the left, no change of any great consequence results from what occurs on Choice 1 In whichever direction the choice is made, the rat must finally go around to the right after Choice 1. Now, on Choice 2 the rat is as likely to turn in one direction as in the other. This is understandable on the basis of the assumption that nothing of any great consequence for learning has occurred on Choice 1, and from the fact that the rat has traveled in a very slow and zig-zag fashion around the right side of the maze and up the central alley, just previous to making Choice 2 This slow movement would nullify the tendency to alternate (if there be such at this time). But whether the rat turns to the left or to the right on the second choice, he must finally go to the left, where he is allowed to escape from the water and maze. Responses in the direction of the goal now have an opportunity to become conditioned to goal and pre-goal stimuli.

Now, after the first, or after a very few trials, all of the rats become very strongly oriented toward the left side of the maze, and both choices thereafter are to the left for many consecutive trials This is shown in Table 2 Table 5, which shows the percentage of each type of the six possible response series, brings out the same point

The rat remains persistently errorless on Choice 2 throughout the training trials. Two important factors no doubt contribute to this great proficiency on Choice 2 First of all, to be sure, is the factor of conditioning. Responses which carry the rat toward the goal are built up to goal and near-goal stimuli, so that these stimuli, whenever presented, call forth these directed responses But a second, and very important factor is the easily acquired habit of simple alternation So it occurs that *the stimuli present to the rat at the choice point plus the retained effects of having just gone to the right* work together to insure that the response on Choice 2 shall always be to the left

The rat, after becoming oriented, also persistently turns left on Choice 1, for many consecutive trials This left-turning habit on Choice 1 is an effect of the conditioning which has occurred on

Choice 2 At the time Choice 1 is made and for several seconds previous to that time, the objective stimulating situation is the same as on Choice 2 This stimulating situation brought forth a left turn on Choice 2, so it now brings forth a similar response on Choice 1. The problem on Choice 1 is similar to that confronting a rat who finds himself in the middle of a long maze the last half of which he has already mastered If the rat is not allowed to retrace, he will be expected to follow the accustomed path Carr and Watson (2) have shown that this is just what occurs. But the *rl* problem requires that the rat turn right on Choice 1, which is requiring him to go opposite the direction he has learned to go on Choice 2 The whole problem for the rat then becomes one of eliminating the left turn on Choice 1 Let us proceed to the method by which this is accomplished

Various types of responses become conditioned into this larger pattern of going left on Choice 1 Anticipatory responses soon extend backward, in this pattern, into the central alley, and then back almost to the starting box. The rat begins hugging the left side of the central alley almost immediately after starting up the alley toward the choice point These responses make it appear to the observer that the rat is trying to turn left before reaching the choice point Simultaneously with most rats anticipatory alley shortening responses begin to appear, and the rat, on many of his trials, does not go to the bottom of the blind alley.

These cul-shortening responses finally extend backward in the pattern of response as far as the choice point and comprise all of the various kinds of responses to which the name "hesitation" has been given.

Hesitation occurs at the choice point because the response to the left is inhibited, and no new pattern of response has been built up to substitute for it The rat cannot go left, its muscles literally "stick" upon arriving at the choice point But it has not learned to go to the right in this or any similar stimulating situation Never before has the rat confronted this portion of the maze at a stand-still Not only must the rat, in order to make a correct response, maintain its posture for inhibiting the left turn, but he must also get a new response under way But upon starting in the right direction, the rat often loses its posture for inhibiting the left response and thus at *last enters the blind alley, most often to the very bottom* It thus occurs that hesitation responses in the early learning trials usually terminate in incorrect responses on Choice 1 As more and more trials are made, and therefore, as more and more hesitation responses occur, the rat becomes more efficient at maintaining its posture for inhibiting the left turns until its hesitation responses bring it into a stimulating situation sufficient to call forth a response to the right

Hesitation responses at the choice point at first occur only occasionally because the cul-shortening response on most trials does not yet extend backward in the error pattern in any large percentage of the trials As the trials accumulate the cul-shortening response occurs on a greater and greater per cent of the trials until hesitation occurs on each trial

Rats 6 and 7 exemplify this shortened cul-response reduced to a vestige All of the rats, on their trips into the blind alley on Choice 1, developed the habit of turning around left before retracing their paths out of the blind alley This habit continued as they shortened their trips into the blind alley A number of rats retained this turn after having eliminated the blind alley. Rats 6 and 7, in 7 5 per cent and 7 2 per cent respectively of their trials after mastery, engaged in this variety of hesitation response The response consisted of turning around left as on a pivot through an arc of 270° and then proceeding to the right

The other types of hesitation responses were also phenomena of the same character as the anticipatory response of cul-shortening The returns back the central alley were examples in which the cul-shortening response extended back beyond the choice point into the central alley. The rat started out from the starting box, swam a part of the length of the central alley, and became inhibited The general behavior of the rat presented the same picture as the cul-shortening responses In the early trials in which this type of response occurred, the rat turned his head toward the left side wall of the central alley, came to a stop, and began retracing back to the starting box. As in the other varieties of hesitation response, so in this one also the rat, after becoming inhibited, became disinhibited in the return back to the starting box, and upon traversing the central alley the second time, turned to the left into the blind alley immediately on Choice 1, without hesitating After some practice in this return type of response, a number of the rats were able to regain one of the more primitive types of cul-shortening responses (stopping in section 6 or 5) after becoming disinhibited by a return to the central alley

The response of swimming straight ahead past the choice point and bumping into the end of the maze also, no doubt, is the result of the inhibiting response of cul-shortening The rat swimming up the central alley becomes inhibited from going left, but has yet learned no other course, so continues by inertia straight ahead into the end wall It is surprising that more cases of this type of response were not present Like the other types of behavior following inhibition, crashing into the end of the maze served the function of disinhibiting the response of rejecting the blind alley So, after recovering from the bump, the rat would turn into the blind alley as if he had not been previously inhibited However, the rat is amenable to conditioning against the effects of such distraction, and after some practice in encountering the end wall is able to maintain the requisite posture for inhibiting entrance into the blind alley, even after engaging in the intermediate activity of the end wall encounter Sooner or later, with most animals, all such hesitation behavior becomes eliminated at the choice point and the rat makes a clean-cut right turn on Choice 1, and the problem is completely solved A later development is that of anticipating a right turn on Choice 1 by hugging the right side of the central alley or by actually turning into the right side of the central alley before arriving at the choice point

In regression these responses may be observed to shift in the opposite direction

CONCLUSIONS

1 All of the rats soon (in from 1 to 8 trials) became oriented toward the goal on both choices of the simple alternation problem (*rl*), and thus developed a blind-alley response on Choice 1

2. Thenceforth the problem of the test was one of eliminating this blind-alley response Twelve of 14 rats eliminated the cul-response, and hence reached the criterion of mastery (three consecutive correct trials), in from 36 to 173 trials The other two failed to master the problem in 200 trials

a The cul-response was eliminated by means of a hierarchy of cul-shortening transition steps With 9 of the rats the cul-shortening process on one or more trials extended back into the central alley, or to the starting box b That transition step in the cul-shortening hierarchy which fell at the choice point resulted in hesitation behavior The early hesitation responses usually ended in failure (i e, entrance into the blind alley). With practice, however, the rats learned to succeed (inhibit entrance into the blind alley) after hesitating at the choice point.

3 Modification of the stimulating situation resulted in a reversion of one or more steps in the hierarchy of types, with rats in all stages of the hierarchy

4 Eight of 12 rats tested were able to extend the learned two-choice simple alternation series, *rl*, to four choices, *rlrl*, in from 40 per cent to 87 per cent of the trials

REFERENCES

1 CARR, H A The Alternation Problem A Preliminary Study *J Anim Behav*, 1917, 7, 365-384

2 CARR, H. A, & WATSON, J W Orientation in the White Rat *J. Comp Neur. & Psychol*, 1908, **18**, 27-44.

3. DUNN, B. M A Comparison of Food Reward and Escape from Water in Motivating Learning in the White Rat. *J Comp Psychol.*, 1935, **19**, 107-112

4 GELLERMANN, L. W The Double Alternation Problem I The Behavior of Monkeys in a Double Alternation Temporal Maze *J Genet Psychol*, 1931, **39**, 50-71

5 ——— The Double Alternation Problem II The Behavior of Children and Human Adults in a Double Alternation Temporal Maze *J Genet Psychol*, 1931, **39**, 197-223

6 ———. The Double Alternation Problem III The Behavior of Monkeys in a Double Alternation Box-Apparatus *J. Genet. Psychol.*, 1931, **39**, 359-389

7 HACK, E. R Learning as a Function of Water Temperature. *J Exper. Psychol.*, 1933, **16**, 442-445.

8. HULL, C. L The Concept of the Habit-Family Hierarchy and Maze Learning *Psychol Rev*, 1934, **41**, 33-54, 134-152

9 HUNTER, W S The Auditory Sensitivity of the White Rat. *J Anim Behav*, 1914, **4**, 215-222.

10 ———. The Delayed Reaction in Animals and Children. *Behav Monog*, 1913, 2 No. 1, Pp v+86

11 ——— Human Behavior Chicago Univ Chicago Press, 1928

12. ——— Kinaesthetic Sensory Processes in the White Rat *Psychol Bull.*, 1918, **15**, 36-37.

13 ———. The Sensory Control of the Maze Habit in the White Rat *J Genet Psychol*, 1929, **36**, 505-537

14 ——— The Temporal Maze and Kinaesthetic Sensory Processes in the White Rat *Psychobiol*, 1920, **2**, 1-18

15 HUSBAND, R W. A Comparison of Human Adults and White Rats in Maze Learning *J Comp Psychol*, 1929, **9**, 361-377.

16 MACFARLANE, D A The Rôle of Kinaesthesis in Maze Learning *Univ Calif Publ Psychol*, 1930, **4**, 277-305

17. MACGILLIVRAY, M E., & STONE, C P The Incentive Value of Food and Escape from Water for Albino Rats Forming the Light Discrimination *J Comp Psychol*, 1931, **11**, 319-324

18 RUCH, F L Food Reward *vs* Escape-from-water as Conditions Motivating Learning in the White Rat. *J Genet Psychol*, 1930, **38**, 127-144

19 WEVER, E G Water Temperature as an Incentive to Swimming Activity in the Rat *J Comp Psychol*, 1932, **14**, 219-224.

Ohio State University
Columbus, Ohio

CONTINUED TRAINING WITH THE OMISSION OF CERTAIN NEBENREIZE*[1]

Stanford University

PAUL R. FARNSWORTH

It is a matter of common observation that, when one learns something, items relatively irrelevant to the matter at hand may become tied into the process. Thus the pianist may suffer considerable loss of ability when his piano stool is not at its accustomed height. The student who is prohibited from stroking his face may find it difficult to do his lessons. Observations of this sort have led to the belief that one's surroundings, the Nebenreize, are of considerable importance in learning situations. And there is research (3, 4) which has been taken to indicate that at least on occasion this is true. But the practical issue is that of finding the conditions under which a theory will hold and those under which it does not apply.

It was a consideration of this practical problem which motivated a previous investigation on the part of the present writer (1). In that research he demonstrated that no ill effect could be shown to operate when college examinations were taken in rooms in which the classes had not been held. Either the Nebenreize here were unimportant or the examination procedures were too crude to show the real effects. This same practical problem motivated the present experiments.

In the present study,[2] groups of college students were trained on the Koerth pursuit rotor[3] and the Dearborn-Coover code test[4] in the

---

*Received in the Editorial Office on October 20, 1936.

[1] The writer gratefully acknowledges the financial aid of the Laura Spelman Foundation. He wishes to thank the Messrs. William Biel and Leonard Ferguson for their technical assistance.

[2] Reported in part at the 1935 meeting of the Western Psychological Association.

[3] The Koerth pursuit rotor (2) sets the task of following a target placed in a phonograph turntable. As employed in the present experiment the motor was electrically driven.

[4] The modified form of the Dearborn code test used in this study included 120 sequences of 5 place numbers. Dr. J. E. Coover, who has been using this test for many years, so modified the original list that (*a*) each digit appears about equally often, (*b*) the digits precede and succeed each other about equally often, (*c*) no digit occurs twice in any sequence of 5, (*d*) no digit falls directly below itself in the columns.

TABLE 1

Dearborn-Coover Code Test

Trials	3-day N=50 Control (M=25) (F=25)	Experimental N=50 (M=50) (F=25)			2-day Control (N=100) (M=50) (F=50)	Experimental (N=50) (M=25) (F=25)		
1st day	Mean	Mean	Diff	C. R.	Mean	Mean	Diff	C R.
1	n 55 64	n 54 08	—1 56	722	n 54 86	n 57 30	+2 44	1 544
2	n 63 98	n 60 08	—3 90	1 970	n 62 03	n 64 22	+2 19	1 147
3.	n 68.90	n 66 32	—2 58	1.162	n 67 61	n 70 22	+2 61	1 210
4	n 72 66	n 70 78	—1 88	681	n 71 72	n 72 78	+1 06	457
5	n 75 70	n 75 76	+0 06	022	n 75 73	n 77 00	+1 27	521
2nd day								
6	n 80 68	n 79 04	—1 64	509	n 79 86	78 58	—1 28	451
7.	n 87.24	n 85 44	—1 80	499	n 86 34	85 42	—0 92	292
8	n 90 98	n 87 94	—3 04	809	n 89.46	87 70	—1 76	573
9	n 92 46	n 88 12	—4 34	1 080	n 90 29	87 86	—2 43	784
10	n 96 46	n 89 26	—7.20	1 805	n 92.86	90.36	—2 50	847
3rd day								
11	n 101 68	93 62	—8 06	1 733				
12	n 104 60	100 40	—4.20	862				
13	n 106 46	101 16	—5 30	1 079				
14	n 106 54	101 86	—4.68	923				
15	n 107 10	102 60	—4.50	872				

C R = Diff / σ_D (Critical ratio)

n = noise present, at all other times it was absent

— = experimental group made poorer score

TABLE 2

KOERTH PURSUIT ROTOR

Trials N=50	3-day Control (M=25) (F=25)	3-day Experimental N=50 (M=25) (F=25)			2-day Control N=100 (M=50) (F=50)	2-day Experimental N=50 (M=25) (F=25)		
1st day	Mean	Mean	Diff	C R	Mean	Mean	Diff	C R
1	n 16 04	n 13 52	—2 52	863	n 14 78	n 11 70	—3 08	381
2	n 23 76	n 22 48	—1 28	347	n 23 12	n 20 92	—2 20	743
3	n 34 28	n 31 86	—2 46	549	n 33 07	n 30 18	—2 89	807
4	n 42 68	n 37 28	—5 40	954	n 39 98	n 38 20	—1 78	396
5	n 48 18	n 45 62	—2 56	502	n 46 90	n +3 18	—3 72	969
2nd day								
6	n 47 60	n 53 66	+6 06	1 242	n 50 63	52 08	+1 45	396
7	n 56 72	n 55 66	—1 06	207	n 56 19	57 40	+1 21	284
8	n 68.58	n 65 90	—2 68	459	n 67.24	65 12	—2 12	435
9	n 70 18	n 70 82	+0 64	114	n 70 50	72 50	+2 00	427
10	n 80 12	n 76 42	—3 70	549	n 78 27	76 16	—2 11	407
3rd day								
11	n 77 40	71 60	—5 80	997				
12	n 86 00	77 80	—8.20	1 380				
13	n 89 26	83 82	—5 44	907				
14	n 90 98	89 36	—1 62	253				
15	n 97 08	97 28	+0 20	034				

presence of a sound of approximately 70 decibels of loudness.[5] One group practiced for one session and the other for two. Following this, the control groups met for another session with the sound present while the experimental groups proceeded in relative silence. Each session the subjects were given 5 two-minute trials on the Dearborn-Coover alternating with 5 twenty-second trials on the Koerth. For the three-day group, then, there were fifteen trials and for the two-day group ten. Both two and three-day sessions were run to tap the learning at different stages. The directions were those traditionally employed for such tests.

The subjects were college sophomores (and a very few juniors) taken from the elementary class in psychology where a requirement to

TABLE 3

3 DAY DEARBORN-COOVER—DATA FOR LAST DAY

Trial	1	2	3	4	5
	Paired for both 1st and 2nd Day within ±4 pts. N=23 (46)				
D	—4.26	—.22	—5.04	+2.09	+1.26
CR	1.58	.09	1.54	.58	.41
	Paired for last 2 trials, 2nd Day within ±2 pts. N=33 (66)				
D	—.15	+3.52	+.42	+4.67	+5.06
CR	.06	1.55	.18	2.11	1.97
	Paired for last trial, 2nd Day within ±2 pts. N=20 (40)				
D	+3.35	+4.15	+6.05	+10.40	+10.50
CR	1.22	1.60	2.18	4.30	3.76
	2 DAY DEARBORN-COOVER—DATA FOR LAST DAY				
	Paired for 1st Day within ±2 pts. N=45 (90)				
D	—3.16	—1.56	—2.71	—2.98	—4.93
CR	1.77	.66	1.18	1.27	1.77
	Paired for last 2 trials, 1st Day within ±2 pts. N=47 (94)				
D	—.87	—.28	—.83	—1.66	—2.26
CR	.53	.15	.44	.92	1.08
	Paired for last trial, 1st Day within ±2 pts. N=38 (76)				
D	—1.05	—1.61	—2.87	—4.03	—4.37
CR	.59	.66	1.29	1.62	1.90

Control heard sound throughout.
— = experimental group made poorer score.
N = Number of pairs.

---

[5] The noise was obtained by the use of a Duo-Art player piano roll which was perforated so as to cause a chord (c-e-g at the level of c1) to be struck once each second.

serve as a laboratory subject for several hours during the quarter has been in force for a number of years. Each of the groups reported on in Tables 1 and 2 was composed of 25 of each sex. An exception was made in the case of the two-day control group which was twice as large as any of the others; its data were derived from those of the first two days and included both the three-day control and the three-day experimental. The designation of subjects as control or experimental was done on the basis of random selection.

An examination of the data of Tables 1 and 2 shows that the control and experimental groups were of roughly equal abilities during the first session. And even when the noise was eliminated for the experimental groups their abilities were still similar to those of the controls. The Koerth and the Dearborn-Coover data presented quite similar pictures. Apparently, then, in these situations the change in Nebenreize was not important.

It should be recalled that the two sets of subjects who featured in Tables 1 and 2 were not individually paired; they were chosen by

TABLE 4

3 DAY KOERTH—DATA FOR LAST DAY

Trial	*1*	*2*	*3*	*4*	*5*
	Paired for both 1st and 2nd Day within ±4 pts N=15 (30)				
D	—10 93	—12 53	—9 07	—10 40	+2 40
CR	1 85	1 32	1 19	1 20	31
	Paired for last 2 trials, 2nd Day within ±3 pts N=35 (70)				
D	—6 40	—12 37	—10 94	—6 69	—2 14
CR	1 49	3 40	3 11	1,48	56
	Paired for last trial, 2nd Day within ±2 pts N=24 (48)				
D	—4.00	—5 50	—3 25	+4 00	+5 83
CR	1 10	1 86	1 11	94	1 65
	2 DAY KOERTH—DATA FOR LAST DAY				
	Paired for 1st Day within ±2 pts N=39 (78)				
D	+2 92	+7 03	+4 31	+7 69	+2 56
CR	77	1 71	93	1 68	52
	Paired for last 2 trials, 1st Day within ±2 pts N=41 (82)				
D	—1 76	+ 29	—1 05	—1 95	—3 34
CR	53	08	28	49	80
	Paired for last trial, 1st Day within ±2 pts N=37 (74)				
D	+1 24	+3 43	+ .19	—2 29	— 70
CR	29	84	04	46	13

random selection. To study the effects of individual pairing, a number of systems were employed (see Tables 3 and 4) The present writer holds no special brief for any one of them To pair score for score was quite obviously out of the question as the available population was too small To pair so that the individual matchings did not deviate more than ±2 points in some groups and ±4 points in another seemed a reasonable procedure, one which yielded from 15 to 47 pairs In certain instances, of course, the individual matchings were perfect

An examination of the data of Table 3 shows only two critical ratios whose magnitudes are 3 or over and these are in the plus direction, i e , the experimental group made the better scores In Table 4 there are again only two critical ratios which are very significant In the main, the values seem to fluctuate in a random manner

RESUME

Groups of students were trained on the Koerth pursuit rotor and the Dearborn-Coover code test for five (or ten) trials in the presence of a sound of about 70 decibels of loudness During the following five trials one-half of the subjects heard the noise and one-half did not No significant nor consistent differences in the performances of the two groups appeared It would follow, then, that the change in Nebenreize was not very effective although it is possible that the elimination of the noise might have caused an appreciable decrement in score had not the noise previously served as a source of distraction. That is, it is conceivable that a true decrease in score due to the change in Nebenreize was masked by an increase in score due to the elimination of a major source of distraction Be this as it may, these data do indicate that the elimination of prominent Nebenreize is not always harmful to one's skills

REFERENCES

1 FARNSWORTH, P R Examinations in familiar and unfamiliar surroundings *J Soc Psychol*, 1934, **5**, 128-129

2. KOERTH, W Eye-hand coordination *Psychol Monog*, 1922, **31**, 288-292

3 SMITH, S, & GUTHRIE, E R General psychology in terms of behavior New York Appleton, 1921. Pp xii+270

4 WILSON, W R The facilitating effect upon learning of accompanying conditioning stimuli Master's Thesis, Univ Washington, 1920

Stanford University
California

A STUDY OF THE CORNELL-COXE PERFORMANCE ABILITY SCALE WITH SUPERIOR CHILDREN*

University of Pittsburgh

HAZEL LINCOLN

The psychological examination has become a valuable aid in school administration Prognosis of success in school adjustment and analysis of existing maladjustments are the work of the school psychologist His basic tool has been the standardized intelligence test, either a group pencil and paper test or the Stanford-Binet test Within limits, these have been useful for predictive and diagnostic purposes, but with changing concepts of intelligence the need has been felt for non-language tests to supplement the usual language type

Studies have shown that wide discrepancies exist between the performance of children with low intelligence quotients on non-language tests and their performance on language tests, discrepancies far beyond what we would be led to expect from correlations between these two types of tests However, no research has been published that would show us what we might expect from superior children in this regard Are there also wide discrepancies between the performance on language and non-language types of tests of superior children, and if so, what do they mean?

This question is of especial importance in the administration of a progressive school where the population is on the whole of superior intelligence Since stress is laid on development of the whole personality of the child and on concrete expression of ideas as well as on the acquisition of academic skills, prognosis of success must take into consideration not only ability to succeed in school work of the academic sort, but also ability to abstract from experiences, to perceive relationships, and to deal with concrete materials, irrespective of language ability.

The problem of this research was one of determining the suitability of the Cornell-Coxe Performance Ability Scale as a performance measure for the guidance of superior children in a progressive school The study developed along two lines, (*a*) the comparison of the scores

---

*Received in the Editorial Office on October 20, 1936

of superior children on the Cornell-Coxe Scale and on the Stanford-Binet test and (*b*) an attempt to determine the significance of the discrepancies that were found in some of the cases.

The subjects for this study were all children with Stanford-Binet intelligence quotients of 110 and above in attendance at a private school The number of subjects was 55. This was approximately 50 *per cent of the school's population.*

The range of the chronological ages of the pupils studied was from 5 years, 6 months to 12 years, 2 months. The median age was 8 years, 5 months The children were enrolled in all grades in the school from kindergarten through sixth grade. Each grade was represented by about the same number of boys as girls, likewise the total study included about the same number of boys as girls

The range of the Stanford-Binet intelligence quotients of the children selected for this study was from 110 to 168 The median intelligence quotient was 121 96 and the semi-interquartile range was 7.39

The Stanford-Binet Revision is administered *in this school* to the child while in the first grade or soon after he enters school The *IQ*'s used in this research were obtained from the files and are the results of tests given within the last five years Twenty-seven of the Binet tests, or just about one-half, were given within the last year and one half From time to time, as a part of the work of the course in intelligence testing, the children in this school have been retested by students. The majority of the children of this study have been retested in this manner, and in none of the cases has any significant difference in *IQ* been found The Binet tests used in this study were given by experienced examiners, 44 of them by the same person

The Cornell-Coxe Performance Ability Scale was administered to 55 children It was impossible to obtain mental ages for four of the subjects from the table of norms given by the authors because the tables do not extend far enough The highest mental age given on the table in 16 years, 8 months, the age norm for a score of 260 on the scale The four subjects whose scores in terms of mental age were not obtainable, had weighted scores as follows. 264, 279, 310, and 316 Since no provisions are given for extrapolation, it can only be noted that the mental ages for these particular subjects is somewhat above 16 years, 8 months Obviously these cases cannot be included in the statistical treatment of the total data, and, it must be borne in mind, that the statistical results given are those of just 51 of the cases, the four cases with the highest weighted scores being excluded.

In order to have comparable scores, the Performance Intelligence Quotient for each child was found by dividing the mental age found on the Cornell-Coxe test by his chronological age. This will be referred to as the P.I.Q. in accordance with the practice of the authors. The range of P.I.Q.'s (51 cases) was 90 to 165, the mean, 128.98, and the median, 130.16. The standard deviation was found to be 16.90. The product moment correlation between the Intelligence Quotients found on the Binet Simon test and the P.I.Q.'s for 51 cases is .378±.08. The authors (1) found a correlation of .79. Very little may be inferred from our coefficient inasmuch as the probable error is so large.

On some cases there was great divergence between the P.I.Q. and the Binet *IQ*. For convenience sake, when the P.I.Q. is higher than the Binet *IQ*, the deviation is called a plus deviation, and the subjects showing such scores are called plus deviates. Where the Binet *IQ* is higher, the deviation is called a minus deviation, and the subjects, minus deviates. The range of deviations is from a minus deviation of 26 points to a plus deviation of 46 points. The number of plus deviations (P.I.Q. greater than Binet *IQ*) was 31, with a mean of 16.3. The number of minus deviations (Binet *IQ* greater) was 18, with a mean of 9.0.

Inasmuch as the table of norms for boys and girls separately shows that, on the whole, girls score lower on this scale than do boys, P.I.Q.'s were figured for each subject from the table of separate norms. The product moment coefficient correlation between the P.I.Q.'s from the separate norms and the Binet *IQ*'s was found to be .313±.09. This coefficient was figured on 47 cases because the highest scores are 248 for boys and 238 for girls, both equal to a mental age of 14 years, 0 months. The four scores omitted in this calculation were 251, 254, 258, and 260.

The second part of this study was undertaken in an effort to explain the significance of wide discrepancies between the P.I.Q. and Binet *IQ* of individual children. It was thought that these discrepancies might be indicative of the presence or absence of special abilities. That these two tests measure to some extent different abilities is shown by the coefficient of alienation of the correlation between the Binet and the Cornell-Coxe, found by the authors of the scale to be .607. Since one of the objectives of this study was to determine the suitability of this scale as a clinical measure for the guidance of superior children, it was felt that if an explanation of these discrepancies could be made, it would be of utmost importance

All individuals showing deviations of more than ten points of P.I Q either above or below the Binet *IQ* were selected for study This criterion of selection is arbitrary and includes 28 individuals, or about one-half of the original number studied. Seven of these were minus deviates, 21 were plus deviates Three lines of inquiry were followed in an effort to determine the presence of special abilities or traits In the first place, since each test of the Cornell-Coxe measures a specific ability, the score of each test was compared with a calculated average score. Second, Stanford Achievement Test scores were compared with Binet *IQ's* and P.I Q.'s for the group of plus deviates and for the group of minus deviates And third, each child was evaluated on a check-list of traits with reference to the direction of deviation

ANALYSIS OF SEPARATE TESTS OF THE CORNELL-COXE SCALE

In order to get an indication of the relationship of the scores on the separate tests of the Cornell-Coxe scale to the total score, an average weighted score was obtained by dividing the total weighted score by six, the number of tests, and comparing each separate weighted score with this average score Then the percentage of each group (plus deviates and minus deviates) varying above and the percentage varying below the average score were computed for each test The differences between these percentages were then compared with the standard errors of the differences.

It would be well at this point to mention the specific abilities which each of the separate tests of the scale attempts to measure The Manikin-Profile Test, the authors quote from Pintner-Paterson, presupposes on the part of the subject an ability to read a diagram, to perceive the seven pieces of the profile as a head, to synthesize the various parts into a whole The Digit-Symbol Test affords a measure of a certain routine kind of learning ability The discussion of the ability measured by the Memory-for-Designs Test is taken from Terman The authors quote as follows:

> Without some analysis it would hardly be possible to remember the designs at all The memory span for unrelated objects is far too limited to permit us to grasp and retain that number of unrelated impressions Success is possible only by grouping the lines according to their relationships, so that

> several of them are given a unitary value and are remembered as one.

The Cube-Construction Test is a test of mental imagery The Picture-Completion Test measures apperceptions, eliminating the factors of training and of special ability or disability in language The Block-Designs Test (2) requires

> the breaking up of each design presented into logical units and a reasoned manipulation of the blocks to reconstruct the original design The results of this activity yield a fair index to this analytic-synthetic power which we have termed "intelligence"

The comparisons between the separate tests and the average scores are given in Table 1. In order to clarify this table, the following example is given The first line of data may be interpreted to read that on the Manikin Profile Test, 70 per cent of the group of plus deviates made scores higher than their calculated average scores, 30 per cent scored lower than their average scores Of the minus deviates 43 per cent scored higher and 57 per cent scored lower. The difference between the two groups is 27 per cent The standard error of the difference is 21.4, and the difference divided by the standard error of the difference is 1 26

TABLE 1

A COMPARISON OF THE SEPARATE TESTS ON THE CORNELL-COXE, SHOWING THE PERCENTAGE OF EACH GROUP SCORING BETTER THAN THE AVERAGE SCORE, AND THE PERCENTAGE SCORING LOWER

	Plus Deviates		Minus Deviates				
Test	% Higher than Av	% Lower than Av	% Higher than Av	% Lower than Av	Diff	σ_{diff}	Diff σ_{diff}
M P	70	30	43	57	27	21.4	1.26
B D	65	35	29	71	36	20.0	1.8
D S	5	95	43	57	38	19.6	1.98
M D.	20	80	29	71	9	19.2	.47
C C	37	63	43	57	6	21.9	.27
P C	47	53	14	86	33	17.0	1.94

From the table it may be concluded that none of the separate tests contribute significantly more or less to the total score in either the group of plus deviates or in the group of minus deviates, since in no case is the difference greater than three times the standard error of the difference Three of the tests show fairly significant differences These tests were the Block-Designs Test, the Digit-Symbol Test, and the Picture Completion Test. From this it might be concluded that the plus deviates do relatively better on the Block Designs Test than they do on the other tests, and consequently that they excel in analytic-synthetic power The plus deviates do relatively poorer on the Digit-Symbol Test, this might indicate that their other abilities surpass their routine learning ability The minus deviates do relatively poorer on the Block-Designs Test, possibly they lack analytic-synthetic power in comparison with their other abilities The minus deviates also scored relatively lower on the Picture Completion Test, a test of apperceptions, than they did on their other tests

COMPARISON OF ACHIEVEMENT TEST SCORES OF THE TWO GROUPS

The New Stanford Achievement Tests are administered as a part of the regular work in this school in grades three through six Scores are obtainable for 16 of our 28 cases of the second part of our study Of these, three were minus deviates, and 13 were plus deviates Little could be done statistically with these scores because the number of cases was small Correlations between the Stanford Educational Quotient, obtained by dividing the educational age, read from the norms, by the chronological age, and the Binet *IQ* was found by Spearman's formula for rank difference method to be .78+, $N=16$ Correlation between the Stanford Educational Quotient and P.I.Q was found by the same method to be 327 Examination of the data showed four cases in which the Educational Quotient was greater than the Binet *IQ* Two of these cases were in the group of plus deviates It will be remembered that the total number in each of these groups was three and 13 respectively

ANALYSIS OF THE TWO GROUPS BY MEANS OF A CHECK LIST OF TRAITS

During the administration of the Cornell-Coxe tests, differences in personality and in methods of attacking problems were noted It was thought that possibly these might be of assistance in explaining the wide divergence between the two tests A list of these personality traits was compiled To these were added several statements concerning the child's handwork, his work in dancing and rhythms, his language ability, and his physical traits This made a list of 17 paired traits This list was submitted to the teacher in an interview Terms as they appeared on the list were defined and explained during the interview and interest elicited. "Attention" and "persistence" were limited to mean quality of attention and persistence in academic work. The following comments were made about three of the children: "attention poor because she is slow in following the class discussion", "J. is persistent and attentive when genuinely interested, as in things involving reasoning", "attention good as long as he is doing something with his hands." These three children were checked as having poor attention, good attention, and poor attention, respectively. The pair of traits designated as "impulsive-stolid" was found not to apply to many of the children, so in many of the cases, it was necessary to disregard it. The teacher did not know, at the time of the interview, of the results of the Cornell-Coxe testing, and the traits were not checked, where judgment was questionable. Table 2 shows the percentages of each group (plus deviates and minus deviates) that were checked for each trait. It shows also the differences and the standard errors of the differences.

Only one of these paired traits is to be found in our two groups with statistically significant difference. This pair is the one we have called "attention good—attention poor." One hundred per cent of the minus deviates (Binet *IQ* higher than P.I.Q.) were checked as having good attention; whereas only 63 per cent of the plus deviates were checked as having good attention. The difference here is 37 per cent and the standard error of the difference is 11.0. In one other trait, the difference is fairly significant. This trait is the one designated as "socially mature—socially immature." Sixty-eight per cent of the plus deviates were checked as socially mature, only 20 per cent of the minus deviates were so checked.

It was anticipated that the two groups would show some difference in language ability. The authors (1) of the test say, "If the Performance mental age is significantly lower, the examiner should consider the possibility of special verbal ability . . . ." However, the two groups show very little difference in the two paired traits listed under language ability, "expresses ideas well—has trouble expressing ideas" and "talks considerably—talks little." Both of the groups express their ideas well and talk considerably, rather than little. Two explanations of the differences between our findings and those of the authors suggest themselves. Either our check list was too rough to catch special verbal ability, or possibly the children of this study, coming from superior homes and being subjected to an extraordinary

TABLE 2

A COMPARISON OF THE PERCENTAGE OF PLUS DEVIATES AND MINUS DEVIATES CHECKED BY THE TEACHER AS EVIDENCING EACH OF THE TRAITS

	Traits	% Plus Deviates	% Minus Deviates	Diff	σ_{diff}	$\frac{\text{Diff}}{\sigma_{diff}}$
PERSONALITY TRAITS						
1	aggressive	71	57	14	21 4	65
	submissive	29	43			
2	socially mature	68	20	48	21 1	2 2
	immature	32	80			
3	persistent	70	80	10	20 6	49
	not persistent	30	20			
4	attention good	63	100	37	11 0	3 36
	attention poor	37	0			
5	impulsive	77	60	17	25 0	68
	stolid	23	40			
HANDWORK						
6	coordination good	68	57	11	21 9	50
	coordination poor	32	43			
7	works swiftly	53	20	33	21 6	1 53
	works slowly	47	80			
8	contributes ideas	78	50	28	23 2	1 21
	contributes little	22	50			
9	works without direction	56	57	1	22 4	04
	needs urging	44	43			
10	plans ahead	87	57	30	21 0	1 43
	no apparent plan	13	43			
DANCING AND RHYTHMS						
11	artistic imagination	88	67	21	20 6	1 02
	little imagination	12	33			
12	graceful	94	71	23	18 0	1 28
	awkward	6	29			

TABLE 2 (continued)

Traits			% Plus Deviates	% Minus Deviates	Diff.	σ_{diff}	$\frac{\text{Diff}}{\sigma_{diff}}$
LANGUAGE ABILITY							
	13	expresses ideas well	68	71	3	20 2	14
		has trouble expressing ideas	32	29			
	14	talks considerably	72	67	5	21 9	23
		talks little	28	33			
PHYSICAL TRAITS							
	15	general health good	87	86	1	15 8	06
		health poor	13	14			
	16	right-handed	95	86	9	13 9	65
		left-handed	5	14			
	17	eyesight good	90	100	10	7 0	1 43
		eyesight poor	10	0			

amount of discussion-teaching in a progressive school, are all possessed of special verbal ability in comparison with an average child

CONCLUSIONS

1 The P I Q's for a group of superior children average higher than do the Binet *IQ*'s for the same group. The variability of the P I Q's is also greater than that of the Binet *IQ*'s

2 The coefficient of correlation between the two tests with this group of children is 378±.08, indicating very little correlation.

3 In using this test with superior children, there is a danger of "running out at the top" with children 11 years and above

4 Deviations between the two tests scores range from a minus deviation of 26 points to a plus deviation of 46 points. The plus deviations average greater and were more frequent than the minus deviations

5 From the check list, it was determined that a greater percentage of plus deviates than minus deviates were giving poor attention to academic work and that a greater percentage of minus deviates give good attention Plus deviates may also be socially mature, and minus deviates, socially immature, these results, however, were not statistically significant

6 This study points to the possibility that this scale is inadequately standardized at the extremes, and that further research should be instituted to determine this.

REFERENCES

1 CORNELL, F L, & COXE, W W A performance ability scale Examination manual Yonkers-on-Hudson World Book Co, 1934

2 KOHS, C Intelligence measurement New York Macmillan, 1923 Pp xii+312

University of Pittsburgh
Pittsburgh, Pennsylvania

THE MECHANICAL ABILITIES OF SIBLINGS*

University of Nebraska

ELLIS K FRYE

PREVIOUS WORK

Similarity among traits of relatives has seldom been disputed, but the degree of similarity and the causes of it have been debated Many investigations have been made with respect to such traits as educational achievement, intelligence, physical measurements, and others

Thorndike (27) reported resemblances of siblings in the abilities as measured by the Institute of Educational Research Tests of Selective and Relational Thinking, Generalization, and Organization as 60 after correction for attenuation Hart (8), using the Army Alpha, National, and Binet intelligence tests found a correlation between pairs of siblings of 447 McNemar (17) gave motor tests to 46 pairs of fraternal boy twins and 47 pairs of identical boy twins In all tests, the correlations were higher for identical twins than for fraternal twins Holzinger's (11) data regarding intelligence agree with McNemar's with respect to the greater similarity between identical than fraternal twins Anderson (1), while not entirely *convinced*, infers that mechanical ability is in large measure an inherited one May (18) studied deception in siblings He found all correlations to be positive, the majority ranging from 20 to 40 Starch's (20) work dealt with the achievement of siblings in elementary school subjects The lowest correlation was 06 in writing, the highest, 49 in reading Kramer and Lauterbach (15) investigated similarity of quality of handwriting between the members of sibling pairs Their highest correlation was 27, their lowest, 09

The above studies have shown rather moderate degrees of similarity between siblings None however dealt with strictly mechanical tests

AGE AND SEX DIFFERENCES

Husband and Ludden (13) concluded from their study of motor skills among men and women that a practical equality exists between the sexes Johnson (14) found only 12 6 per cent of the women ex-

---

*Received in the Editorial Office on October 21, 1936

ceeding the average of the men in tight wire walking. He thought this might be due to the fact that the women showed more fear. Bryan (2) reported insignificant mean differences between boys and girls in precision Lehman and Witty (16) made lists of activities of boys and girls. The boys included 37 of mechanical and motor nature; the girls 16 Wells (28) found that women became more fatigued by tapping than men and also that they were more responsive to affective elements. Young's (30) data portray boys slightly superior in pattern-seeking ability, girls in skill. Fisher (6) decided that pubescent factors must be at work, for girls reached their maximum speed at thirteen and boys at fourteen in inserting numbers in geometrical forms Both declined temporarily after that age. Whitman (29) also reported a slight decrease in rate at the age of fourteen. His subjects removed nuts from bolts Schultz (23) stated that boys were definitely superior to girls in a mechanical feeding machine test and that the scores increased directly with age. Perry's (22) data give evidence of boys' superiority in responding to the "given" steps in geometry but of girls' superiority in "link" steps Spearman (25), after reviewing a number of investigations concerning sex differences of a somewhat mechanical nature, concluded that boys are distinctly superior to girls in the reducing of spatial relations, but that the evidence of this difference being really innate is still dubious

NATURE OF MOTOR ABILITIES

Seashore (24) has published a number of articles dealing with motor skills These skills consist of such units as. eye-hand coordination, speed of manipulation, precision, and speed in discrimination reactions. The average correlations among such skills was .25. Patterson and Elliott (20) as a result of a rather comprehensive study of the nature of mechanical ability concluded, among other things, that it probably does not involve any single factor and that it is unique with respect to both intelligence and motor ability. Harvey (9) in reviewing the work of Cox, a student of Spearman's, has pointed out that Cox's conclusion agrees with Patterson's and Elliott's in that both isolated *"M"*, but that the latter demonstrated it to be made up of a number of highly specific factors while the former demonstrated it to be a single group factor. Muscio (19) emphasized the idea that there is no such thing as a "motor type," that motor capacities are relatively independent of intelligence and that from the point of view of vocational guidance every occupation will require specific vocational tests in the selection of workers Farmer (4) administered what he termed aestheto-kinetic coordination tests among which he found an average correlation of 25 Perrin (21) secured low positive correlations among a group of 17 motor tests He decided that motor ability is not general but rather definitely specialized and that a complex motor performance cannot be explained in terms of a number of unit motor functions Garfiel (7) got correlations comparable with Perrin's but attached more significance to the low positive correlations than Perrin did Carter (3), using the tetrad-difference method, found no general factor of mechanical ability but rather a number of group factors which overlap

THE PROBLEM OF THIS STUDY

The specific problem for consideration in this study is to ascertain what relationship exists between the mechanical abilities of siblings. The term "mechanical" as used here pertains to the ability to discern relationships among objects as well as to perform muscular operations. The term "abilities" was used rather than "aptitudes" because the latter is quite generally interpreted as though hereditary traits were involved. The term "abilities" is quite free from such interpretation as well as those relating to the acquired traits The problem is distinctly not one of accounting for any relationship that is found between the mechanical abilities of siblings.

Subsidiary problems given consideration are: Are there sex differences? Are there age differences? What is the nature of mechanical abilities?

In order to ascertain the relationship between siblings with respect to mechanical abilities, three tests of mechanical abilities were given to 200 siblings. Subjects were drawn from high schools in East Central Nebraska

The following tests were given · Stenquist Assembly Test, Series 1; Detroit Manual Ability Tasks, prepared by Alexander C Crockett; and the Wiggly block, designed by Johnson O'Connor.

The Stenquist Test required the subject to assemble properly ten different articles such as, a cupboard catch, bicycle bell, push button. Thirty minutes are allowed for this work with a bonus for doing it in less time

The Manual Ability Tasks consist of three parts, the first, requires the subject to screw nuts on bolts, three trials, each of 2 min.

duration, the second, requires the packing of 1 in. cubical blocks in a container approximately the size of a cigar box, three trials, each of 1 min duration, the third, a precision test in which the subject places the blocks on a narrow board, four trials, each of 1 min duration

The O'Connor Wiggly block is composed of nine blocks, each 9 in long These blocks have irregular faces, but when properly put together form a large block with six plain sides The subject's task is to assemble the block three times, as quickly as possible

Results

One rather simple method of determining whether members of the same family resemble each other more or less than any two members of a heterogeneous population is to count the number of times the members of the same family vary in the same or opposite direction from the mean In the long run, it would seem that pairs selected at random would both have scores above the mean as often as they would have scores below Means were calculated separately for each age and sex Since 12 of the families contained three siblings and one family four siblings, the total number of pair combinations was increased. In a few instances a subject did not take a particular test. These were due to interruptions. Table 1 shows that the members of sibling pairs varied more often in the same direction from the mean than in the opposite direction in the case of every test except the Wiggly block Here there is a difference of one in the opposite direction.

TABLE 1

Number of Times Members of the Same Family Varied in the Same or Opposite Directions from the Means of Their Respective Ages

Wiggly Block		Bolts		Packing		Precision		Stenquist	
S	O	S	O	S	O	S	O	S	O
58	59	67	50	79	38	63	54	74	45

The sum of all pairs varying in the same direction is 351, of those in the opposite direction, 246 Thus there is some indication of resemblance, and the computation of reliability measures is suggested

The second method used in ascertaining the resemblance of siblings was the double correlation method Each individual appears on both the X and Y axis of the scatter diagram, making both the plots symmetrical, with N equal to the number of entries. Where there were more than two siblings in a family, all *combinations were entered.* If there were three in a family, then siblings 1 and 2, 1 and 3, and 2 and 3 were entered. Since the double method was used, there was a total of six entered for a family of three

The single interclass method of correlation might have been used, providing each family of siblings had been listed alphabetically This would have controlled the age factor Again, the single method might have been used had each family of siblings been listed according to age. The age factor could then have been determined by the partial correlation technique

However, the interclass, double correlation, method appears to be the best R A. Fisher (5, p 180) says, "the interclass correlation is not an estimate equivalent to an interclass correlation, but is some*what more accurate* "

The formula, $r = 1 - \dfrac{\sigma^2 d}{2\sigma^2}$ was used for finding the correlation. This formula facilitated the work of correlation and allowed the resemblance to be expressed in percentage It is derived from the basic fact that correlations may be found either from *products* of deviations or from *differences*

The correlations obtained between the siblings in each test are given in Table 2

TABLE 2

STATISTICAL RESULTS OF SIBLING STUDY

(1) Tests	(2) Pairs	(3) r	(4) σ	(5) σd	(6) σdo	(7) 6-5	(8) Pct of 7 to 6
W B							
Boys	25	—18 ± 11	2,726	4 189	3 854	—334	—8
Girls	26	,12 ± 11	3 591	4 753	5 078	324	6
Bolts	121	13 ± 05	9,480	12 334	13 405	1 071	9
Pack	119	52 ± 04	21,970	21 466	31 066	9 600	31
Prec	120	23 ± 04	19 590	23 980	27 700	3 720	13
Stenq.							
Boys	34	19 ± 09	8 000	10 209	11 312	1 103	11
Girls	32	50 ± 09	8 570	8,585	12,118	3 532	29

Since there was considerable indication of sex differences in the Wiggly block and significant sex differences in the Stenquist test the sexes were correlated separately.

Column 1 *contains* the *names* of the tests Column 2 shows the number of pairs of siblings Column 3 gives the correlations Column 4 gives the standard deviations of the scores of all cases. Column 5 gives the standard deviations of the differences between the paired scores Since the double correlation method was used, these differences were taken as both plus and minus. Column 6 gives what the sigmas of the scores would be if the pairs were selected at random and the resulting correlations were zero Column 7 gives the difference between the sigma of the differences between the siblings and the sigma of the differences between pairs with a correlation of zero Column 8 shows the per cent of Column 7 to 6.

An inspection of Table 2, Column 3, shows all positive correlations except in the case of the Wiggly block test for boys. The correlations for packing, precision, and girls' Stenquist Tests are high enough to be significant since each is larger than 4 times it's probable error All per cents, too, in Column 8 are positive except for the same test, Wiggly block for boys. It is evident that the higher the amounts are in these two columns, the closer the resemblance between siblings.

The causes for certain tests giving correlations too low to be significant may be due to the inferior construction of the tests themselves, errors in administering them, or to the possibility that siblings actually do not resemble one another in the traits measured by these tests Still another possibility, that no resemblance actually exists, must be dismissed with the statement that as far as the data of the present study are concerned, no resemblances statistically significant exist However, all correlations, except for the boys Wiggly block, are positive Even here it is positive when subjects who failed to complete the test were taken into consideration

It was not the purpose of this investigation to ascertain why sibling resemblances do, or do not, exist This would immediately raise the nature-nurture controversy. It would be contentious to assert that any sibling resemblance was due entirely to either nature or nurture in that the controls for the latter were not provided.

The results just presented, that is, rather low positive correlations between siblings in mechanical tests, are in fair agreement with the results of other studies regarding the resemblance of siblings in certain achievement tests The correlations are lower than those found in intelligence tests and also lower than those found where twins were used rather than mere siblings.

The only conclusion that is statistically defensible, is that the score of one of a pair of siblings in each of the packing, precision, and girls'

Stenquist tests is predictive of the score of the other one of the pair of siblings in the same test. Since the correlations are only fair, this prediction is by no means infallible.

SEX DIFFERENCES

Two methods of comparing the scores of boys with those of girls were used: first, the per cent of boys reaching or exceeding the median score of the girls; secondly, the ratio of the differences of the means to the sigma of the differences of the means.

Table 3 shows the overlapping of the medians of the boys upon those of the girls. Fifty per cent indicates equality between the sexes, while any per cent higher than that shows superiority of the boys. The *F* columns show the number of boys who equalled or excelled the girls for each of the ages fourteen to nineteen respectively. All per cents except for the fourteen-year-olds were in favor of the boys on the Wiggly block, for bolts, the same condition existed save for eighteen-year-olds, for packing, fifteen-year-old boys equalled the record of girls, but only sixteen-year-old boys exceeded the girls' record; for precision, fourteen-year-old boys exceeded and nineteen-year-old boys equalled the girls, the boys exceeded the girls at all ages on the Stenquist test.

TABLE 3

PER CENT OF BOYS REACHING OR EXCEEDING THE MEDIAN GIRLS' SCORE ON THE WIGGLY BLOCK, DETROIT MANUAL, AND STENQUIST TESTS

Age	Over-lapping	Wiggly Block		Detroit Bolts		Manual Packing		Ability Precision		Stenquist T Score	
		F	%	F	%	F	%	F	%	F	%
14	B – G	6	43	9	64	2	14	8	57	13	93
15	B – G	14	72	14	72	9	50	8	44	15	83
16	B – G	13	68	14	74	15	79	8	42	17	89
17	B – G	11	61	11	61	2	11	6	33	15	82
18	B – G	11	92	5	42	4	33	5	42	12	100
19	B – G	8	67	8	67	5	42	6	50	8	75

A more accurate method of showing the differences between the sexes is that of finding the critical ratio of the differences of the means. In order that critical ratios be significant they must be at least as high as three. The only test in which this situation existed was with respect to the Stenquist test. In this test the critical ratios were higher than three at all ages except nineteen. All favored the boys.

Boys showed some tendency to excel in the Wiggly block and bolt tests, the girls in the packing and precision tests. Many would agree that boys had had more experience in the tests in which they excelled. The tendency for girls to excel in packing and precision seems to justify the rather common practice of employing girls in industries where speed and precision are at a premium.

AGE DIFFERENCES

Critical ratios were also found between the differences of the means of subjects of successive ages. In no case was a critical ratio significant. However the data show a tendency for the subjects to reach their highest efficiency at ages seventeen and eighteen with a slight falling off at nineteen. The only exceptions are for nineteen-year-old girls in the Wiggly block and sixteen-year-old boys in packing.

The ranges of scores for the mechanical tests are given in Table 4.

TABLE 4
RANGES OF SCORES FOR MECHANICAL TESTS

Test	Highest Score	Lowest Score	Range	Highest Score divided by Lowest Score
*Wiggly Block	1.09	15.89	14.80	14.6
Bolt	87	27	60	3.2
Packing	213	71	142	3.0
Precision	173	41	132	4.2
Stenquist (R.S.)	110	35	75	3.1

*Eighteen students failed to complete the test within the time limits, so exact data are not available. Scores are expressed in work units so that the lowest score represents the best performance.

The above data show that the best performance in each test is a little more than three times that of the poorest, except for the Wiggly block where it is approximately fourteen times the poorest.

THE NATURE OF MECHANICAL ABILITY

Much has been written concerning the nature of mechanical ability. Is one justified in speaking of a general mechanical ability or are there many unitary traits that constitute this ability? Correlations among the mechanical tests suggest an answer to this question. Such correlations are presented in Table 5.

All correlations between Wiggly block and other tests are negative

TABLE 5

CORRELATIONS BETWEEN MECHANICAL ABILITY TESTS SCORES FOR GIRLS

	Age	W B	Bolts	Pack	Preci-sion	Sten-quist
Age		— 26	18	— 01	13	07
W B			— 17	— 17	— 35	— 52
Bolts				54	49	58
Pack					47	30
Prec						42

However, they must be interpreted as positive since the score is in units of time for the Wiggly block, but in work units for all other tests

The first line shows the correlations between age and each test The highest correlation is between age and Wiggly block The lowest is between age and packing. All these are so low that they have little or no significance This fact was also shown in a previous section by means of critical ratios

The second line shows low correlations between Wiggly block and bolts and between Wiggly block and packing. Since ability to perceive relationships is essential in the Wiggly block and speed is essential in the other two tests, one would not expect close agreement Speed is not the only important element in the precision test, but no good *reason is apparent for the higher correlation,* — 35, between this test and the Wiggly block. The fair correlation of — 52 between Wiggly block and the Stenquist test might partially be accounted for in that both require considerable ability to see relationships

Line three shows the correlations between bolts and packing, precision, and Stenquist respectively The common element of speed, as well as manipulation has already been mentioned in respect to the bolt and packing tests Here a fair correlation, 54, is found Bolts and precision show a correlation of 49 Some precision is also required for the bolt test; both require manipulation and speed A fair correlation, 58, between bolts and the Stenquist test might possibly be accounted for since each deals with objects quite similar in nature

In the fourth line, packing and precision, both requiring speed and manipulation of the same objects, give a correlation of 47 Between packing and Stenquist the correlation is 30 Here there is manipulation of different objects, speed being stressed in the first test, relationship in the second.

In the fifth line, a fair correlation is found between precision and Stenquist. Some precision is required for the latter test

Rather low correlations, yet all positive with one exception, have been shown to exist among the mechanical ability scores of boys. Practically the same results may be discerned in an examination of Table 6 which presents the results for the girls.

TABLE 6

CORRELATIONS BETWEEN MECHANICAL ABILITY TESTS SCORES FOR GIRLS

	Age	W B.	Bolts	Pack	Precision	Stenquist
Age		—10	34	19	.24	22
W B			—18	—17	—24	—.13
Bolts				.37	37	.20
Pack.					.25	09
Prec						12

All correlations between Wiggly block and other tests are minus, but must be interpreted as plus for the same reason as given above for the boys.

Line 1 shows age as being of little or no significance.

Line 2 shows very much the same correlation between Wiggly block and other tests, with the exception of Stenquist, as was true of the boys. Here the correlation for the girls is —.13 while for the boys it is —.52. The girls' scores in the latter tests were much lower than the boys, but they were consistently lower so that this offers little explanation for the difference in correlation.

Line 3 shows results somewhat lower than those of the boys. This is especially true in regard to the Stenquist test, a situation which was also evident in the previous line

Lines 4 and 5 continue to show the same tendency toward lower correlations for girls than for boys.

These low correlations may be due either to the organization and administration of the tests themselves, or to the nature of the traits tested.

As to the administration of the tests, the writer can merely assert that the utmost care was taken in following the directions for each test given, while concerning the organization and content of the tests, nothing was done save to utilize those tests whose authors are of recognized standing.

The other possible explanation, that of the low correlations being due to the nature of the traits tested, appears to be the most acceptable Hull (12, p 213) says,

> Summarizing the evidence as to the relations among motor activities, we may say that as a rule they correlate very low with each other, but they occasionally show rather high correlations This is likely to be the case when the apparent similarity between the two activities is very marked These facts seem effectually to preclude belief in a two-level theory of aptitudes made up of complex intellectual and simple motor abilities

POSSIBLE PRESENCE OF A GENERAL FACTOR

The fact that the correlations are positive, even though low, is not easily dismissed from one's mind The question arises whether this slight agreement between pairs of tests might not be due to the existence of a general factor

The best method of determining whether a general factor is present is the tetrad difference method. This method was used for the intercorrelations of the boys' as well as the girls' tests

The tests and the intercorrelations were as follows·

	Boys					Girls			
	2	3	4	5		2	3	4	5
1	17	17	35	52	1.	.18	.17	24	.13
2		54	49	58	2		37	.37	20
3			47	30	3			25	09
4				42	4				12

1 Wiggly block
2. Bolts
3 Packing
4 Precision
5. Stenquist

Holzinger's (10) formula was used·

$$t_{1234} = r_{12} \quad r_{34} — r_{13} \quad r_{24} = 0$$

Here t_{1234} is the product of r_{12} (first two subscripts) and r_{34} (second two subscripts) minus the product of r_{13} (first and third subscripts) and r_{24} (second and fourth subscripts)

The probable error formula for r_{1234} is

$$\frac{6745}{\sqrt{N}}—\Big\{ r_{12}^{2} + r_{13}^{2} + r_{24}^{2} + r_{34}^{2} —2 \quad (r_{12}\, r_{13}\, r_{23} + r_{12}\, r_{14}\, r_{24} + r_{13}\, r_{14}\, r_{34} + r_{24}\, r_{34}) + 4\, r_{12}\, r_{13}\, r_{24}\, r_{34} \Big\}^{½}$$

The probable error formulae for all of the tetrads were taken directly from Holzinger (10).

The tetrads, with their probable errors, are given in Table 7

TABLE 7

Tetrads and Probable Errors of Mechanical Tests

Tetrads	Boys	Girls
t_{1234}	—0034 ± 0271	—0179 ± 0246
t_{1243}	—1091 ± 0400	—0438 ± 0173
t_{1342}	—.1057 ± 0413	—0259 ± 0333
t_{1235}	—.0476 ± 0250	—0178 ± 0168
t_{1253}	—2298 ± 0420	—0319 ± 0260
t_{1352}	—1822 ± 0567	—0140 ± 0286
t_{1245}	—1316 ± 0287	—0264 ± 0198
t_{1254}	—1834 ± 0338	—0265 ± 0254
t_{1452}	—0518 ± 0549	—0001 ± 0300
t_{1345}	—0336 ± 0181	—0012 ± 0183
t_{1354}	—1730 ± 0432	—0121 ± 0200
t_{1453}	—1394 ± 0454	—0109 ± 0224
t_{2345}	—0798 ± 0258	—0111 ± 0300
t_{2354}	—0458 ± 0482	—0056 ± 0268
t_{2453}	—1256 ± 0351	—0167 ± 0261

An inspection of the tetrads for boys in Table 7 reveals that all other tetrads than t_{1234}, t_{1235}, t_{1452}, t_{1345}, and t_{2354} are larger than twice their probable errors Thus ten out of fifteen tetrads do not equal zero, consequently, the conditions necessary for proving the T-Factor Theory are not present

In the case of the girls, the proper conditions are present in all tetrads except t_{1243}

Conclusions

The purposes of this investigation were first, to ascertain if siblings resemble each other in mechanical abilities, second, to determine the extent of sex, age, and individual differences in mechanical abilities, third, to attempt to discover the nature of mechanical abilities

Three mechanical ability tests were given to 200 siblings of high-school age

The conclusions to be drawn from this investigation are

1. Using the method of variance from the mean, sibling resemblance was present in all tests except the Wiggly block

2 Positive correlations were found between siblings in all tests except the boys' Wiggly block test Even here, they were positive when the boys who did not complete the test were included In the

case of three tests (the girls' and boys' scores were correlated separately) were the coefficients high enough to be significant

3 When the method of ascertaining the per cent of one sex that excels the other was used, boys were superior in the Wiggly block, bolt, and Stenquist Tests, girls were superior in precision and packing.

4 The Critical Ratios between boys' and girls' scores were significant only in the Stenquist Test Here the boys were definitely superior

5 No significant differences were found between any successive ages in any tests Graphs of mean scores at each age showed strong tendencies for both boys and girls to reach their highest efficiency at ages seventeen and eighteen.

6 Ranges between highest and lowest scores in the mechanical tests were approximately three for all texts except the Wiggly block, here it was fourteen

7 Negligible to fair correlations were found among the mechanical tests.

8 In fourteen out of fifteen tetrads for girls, conditions were present which would substantiate the Two-Factor Theory; in only five out of fifteen tetrads were these conditions true for boys

REFERENCES

1 ANDERSON, L D The relationship of certain environmental factors to measures of mechanical ability *27th Yearbook, Natl Soc Stud Ed, Part 2,* 1928, 137-150

2 BRYAN, W L On the development of voluntary motor control *Amer J Psychol*, 1892, **5**, 125-204

3 CARTER, H D The organization of mechanical intelligence *J Genet Psychol*, 1928, **35**, 270-285

4 FARMER, E A group factor in sensory-motor tests *Brit J. Psychol* (Gen. Sec), 1927, **17**, 327-334

5 FISHER, R A Statistical methods for research workers London Oliver & Boyd, 1930 Pp xiii + 283

6 FISHER, V. E A few notes on age and sex differences in mechanical learning *J Educ Psychol*, 1927, **18**, 562-564

7 GARFIEL, E The measurement of motor ability *Arch of Psychol*, 1923, No 62 Pp 46

8 HART, H Correlations between intelligence quotients of siblings *School & Soc*, 1924, **20**, 381

9 HARVEY, O L Mechanical aptitude or mechanical ability?—a study in method. *J Educ Psychol*, 1931, **22**, 517-522

10 HOLZINGER, K J Statistical resume of the Spearman two-factor theory Chicago Univ Chicago Press, 1930 Pp 43

11. ——— The relative effect of nature and nurture influences on twin differences. *J. Educ. Psychol.*, 1929, **20**, 241-248.

12. HULL, C. L. Aptitude Testing. Yonkers: World Book Co., 1928. Pp. xiv+535.

13. HUSBAND, R. S., & LUDDEN, M. S. Sex differences in motor skill. *J. Exper. Psychol.*, 1931, **14**, 414-422.

14. JOHNSON, G. B. Sex differences in learning to walk a tight wire. *J. Soc. Psychol.*, 1932, **3**, 114-118.

15. KRAMER, K., & LAUTERBACH, C. E. Resemblance in the handwriting of twins and siblings. *J. Educ. Res.*, 1928, **18**, 149-152.

16. LEHMAN, H. C., & WITTY, P. A. Sex differences: interest in tasks requiring mechanical ability and motor skill. *J. Educ. Psychol.*, 1930, **21**, 239-245.

17. MCNEMAR, R. Twin resemblance in motor skills, and the effect of practice thereon. *J. Genet. Psychol.*, 1933, **42**, 70-97.

18. MAY, M. A. Sibling resemblance in deception. *27th Yearbook, Nat'l Soc. Study of Educ., Part 2*, 1928, 161-179.

19. MUSCIO, B. Motor capacity with special reference to vocational guidance. *Brit. J. Psychol.*, 1922, **13**, 157-184.

20. PATTERSON, D. G., & ELLIOTT, R. M. Minnesota mechanical ability tests. Minneapolis: Univ. Minnesota Press, 1930.

21. PERRIN, F. A. C. An experimental study of motor ability. *J. Exper. Psychol.*, 1921, **4**, 24-56.

22. PERRY, W. M. Are boys exceeding girls in geometric learning? *J. Educ. Psychol.*, 1929, **20**, 270-279.

23. SCHULTZ, R. S. A test for motor capacity in the industry and in the school. *J. Appl. Psychol.*, 1928, **12**, 169-189.

24. SEASHORE, R. H. Stanford motor skills unit. *Psychol. Monog.*, 1928, **39**, 51-65.

25. SPEARMAN, C. The abilities of man. London: Macmillan, 1927. Pp. vi + 415.

26. STARCH, D. Inheritance of abilities in school studies. *School & Soc.*, 1915, **2**, 608-610.

27. THORNDIKE, E. L. The resemblance of siblings in intelligence. *27th Yearbook Nat'l Soc. Nature and Nurture, Part I*, 1928.

28. WELLS, F. L. Sex differences in the tapping test. *Amer. J. Psychol.*, 1909, **20**, 353-363.

29. WHITMAN, E. C. A brief test series for manual dexterity. *J. Educ. Psychol.*, 1925, **16**, 118-123.

30. YOUNG, H. H. Slot maze. *Psychol. Clin.*, 1922, **14**, 73-82.

Bradley Polytechnic Institute
Peoria, Illinois

THE EFFECT OF FUNCTIONAL PERIODICITY ON THE LEARNING PROCESS*

The Mills School for Kindergarten-Primary Teachers

ORPHA MAUST LOUGH

More than three years of the most active part of a woman's life is involved in the destructive phase of the menstrual cycle On a minimum basis of three days for each catamenial period, these physiological changes which occur thirteen times yearly, take place during a total of thirty-nine days, or twelve per cent, of each year for approximately half of the life span of the average woman, thus making a total of more than eleven hundred days, or, on the basis of eight hours to the day, approximately, nine thousand working days.

The study of the effects, both mental and physical, of functional periodicity is of greatest importance Now when women are competing more and more with men on a high level of intellectual efficiency, the effect of functional periodicity upon mental ability is a matter not only of educational importance but also of widespread sociological and economic significance.

Although many superstitions have arisen, and though some writers have philosophized upon the mental effects of the gynaecological periods, little or no accurate scientific information has been obtained regarding the psychological aspects of menstruation in an adequate sampling of women of approximately the same age and occupation.

Numerous studies have been made on the physiological aspects of menstruation but the psychological effect of woman's physiological rhythm has been given but scant consideration. The studies which have been made by Hollingworth, Eagleson, Seward, and Sowton and Myers, indicate for the most part that there is no reason to believe that functional periodicity effects considerably or constantly the quality or quantity of woman's work. Ordinarily, menstruation has no noticeable effect on working capacity among normal, healthy women, although there are individual differences and some women

---

*Received in the Editorial Office on November 12, 1936

find an appreciable lowering of efficiency during catamenial periods Certain criticisms might be made of these studies In several of the investigations, the results were based on a meager number of subjects who represented a wide range of ages and were, therefore, not strictly comparable The groups studied were highly selected In some instances the conclusions were based on inadequate data, showed superficial and inadequate statistical treatment of data, or used poor techniques In two studies, men were used as controls The nonmenstrual period was compared with the menstrual period, whereas each phase of the cycle requires separate analysis and comparison with each of the other phases The fluctuations which are found in all learning curves were not given sufficient consideration In some of the experiments the subjects were acquainted with the nature of the investigation and thus opened the results to the influence of suggestion While all of these criticisms do not apply to any one study, they are, nevertheless, indicative of the pioneer character of the work However, each of these studies must be recognized for its contribution to the understanding of the nature of functional periodicity

This investigation was inspired by the great need of accurate, scientific information regarding the psychological effects of woman's fundamental physiological rhythm The subjects in this study were 96 unmarried women students enrolled in the first three years of a college[1] curriculum Their ages ranged from sixteen years, eleven months, to twenty-five years, three months The total number of subjects and the small range of ages would indicate that they constituted an adequate random sampling of women students in the first three years of college The subjects came from families of approximately the same social and economic status

Each subject provided her own control in as much as each of the four phases of her menstrual cycle was compared with the other three phases of her cycle. From the correlations established by these comparisons, it was possible to determine the effect of each phase upon the learning process

The learning test was administered every consecutive school day for a total of 32 days to 65 students, and for a total of 40 days to

---

[1] Students at the Harriette Melissa Mills Training School for Kindergarten and Primary Teachers, 66 Fifth Avenue, New York, N Y

31 students The learning test was the digit-symbol type and involved the solving of simple *standardized equations taken from the Clapp Number Combination Test, Test I* These equations were expressed in non-arithmetical symbols, according to the code supplied at the top of each test form. The daily five-minute testing period avoided fatigue effects and functioned in maintaining sustained interest and attention The quantitative data were delimited to the scores on the five forms of the five-minute learning test, *Forms A* and *B* of the *Henmon-Nelson Test of Mental Ability,* and the *Thurstone Personality Schedule, Clark Revision. Form B* of the *Henmon-Nelson Test of Mental Ability* was given during menstruation and *Form A* between the catamenial periods

Regularity of functional periodicity, pain during menstruation, habits of elimination, and habits of worry as given from memory by each subject on the first day of the experiment, together with the attributes of emotional disturbance, pain, and worry recorded daily by each subject during the entire experiment, constituted the qualitative data

The problem was presented as only an experiment in learning, an experiment that would enable each subject to determine the way in which she herself learned This technique served to keep the subjects naive as to the true nature of the research and to minimize the traditional or psychological effects of menstruation On the first day of the experiment each subject was given a sheet of graph paper on which she was instructed how to keep her own record of the total number of equations she completed each day In this way, each subject was motivated by competition with her own past record

Daily records[2] were kept for each subject, (*a*) the total number of equations completed, (*b*) the number of equations correctly completed, (*c*) the number of equations incorrectly completed From these records was plotted the learning curve for each subject. Examination of these parabolic curves indicated that (*a*) there was learning (incremental change resulting from practice and its attendant conditions), (*b*) the learning tests were valid, (*c*) for each subject there were daily fluctuations in speed and accuracy of learning, (*d*) the speed and accuracy of learning varied widely from subject to subject For many of the subjects there appeared to be a rhythm in performance, but the cause, whether innate, accidental, due to the day of the week, mental set, the menstrual cycle, or other variables, was not clearly indicated by inspection.

---

[2] The data of this research are filed in the library of the School of Education, New York University

The raw scores obtained on the daily learning tests were converted into *T*-scores (a variation of the sigma score or *z*-score) in order to make them comparable from day to day during the entire experiment. The *T*-scores resolved the data of raw scores into an extremely flexible and usable form. They made possible the comparison of the records of each subject, during the four phases of her menstrual cycle, with the records of other subjects during their menstrual cycles, although the actual experimental days on which each of these phases fell might be widely separated.

The *T*-scores on the learning tests were grouped according to the days on which the physiological changes take place in the cycle in order to compare the subject's mental efficiency during the different phases of functional periodicity. Physiologists who have studied the physiological changes taking place during the menstrual cycle, have established four distinct phases: (*A*) the menstrual phase which is the destructive stage, (*B*) the premenstrual phase, (*C*) the postmenstrual phase; (*D*) the intermenstrual phase. Correlations were computed between the average *T*-scores for the equations correctly completed[3] and for equations incorrectly completed on the learning test for each of these four phases, Forms *A* and *B* of the Henmon-Nelson Test of Mental Ability, and the Thurstone Personality Schedule, Clark Revision, as shown in Table 1 and Table 2.

From Table 1 it is apparent that the correlations between the average *T*-scores of the equations correctly completed during each of the four phases of the menstrual cycle were very high, ranging from .94±.007 to .97±.005. In Table 2 the correlations between the average *T*-scores of equations incorrectly completed were high, ranging from .71±.033 to .79±.026, but were not as high as for the equations correctly completed. This difference in the relationships was probably due to the fact that ordinarily the number of equations incorrectly completed was very small, when a few of the

---

[3] Hereafter, the scores on equations correctly completed (right) are designated *A*, menstrual phase, *B*, premenstrual phase, *C*, postmenstrual phase, *D*, intermenstrual phase. The scores on equations incorrectly completed (wrong) are designated *a*, menstrual phase, *b*, premenstrual phase, *c*, postmenstrual phase, *d*, intermenstrual phase.

TABLE 1

CORRELATIONS OF AVERAGE T-SCORES OF EQUATIONS CORRECTLY COMPLETED (RIGHT) DURING THE FOUR PHASES OF THE MENSTRUAL CYCLE, FORMS A AND B OF THE HENMON-NELSON TEST OF MENTAL ABILITY, AND THE THURSTONE PERSONALITY SCHEDULE, CLARK REVISION

A Ave. *T*-score Menstr	*B* Ave *T*-score Premenstr	*C* Ave *T*-score Postmen.	*D* Ave *T*-score Intermen	Henmon- Nelson Form *A*	Henmon- Nelson Form *B*	Thurstone Personal. Schedule
	96± 005	97± 005	95± 007	45± 055	45± 055	—16± 067
		.95± 006	94± 007	46± 054	46± 054	—13± 068
			95± 007	42± 057	43± 056	—17± 067
				43± 056	43± 056	—14± 067
					80± 025	—06± 068
						—15± 067

TABLE 2

CORRELATIONS OF AVERAGE T-SCORES OF EQUATIONS INCORRECTLY COMPLETED (WRONG) DURING THE FOUR PHASES OF THE MENSTRUAL CYCLE, FORMS A AND B OF THE HENMON-NELSON TEST OF MENTAL ABILITY, AND THE THURSTONE PERSONALITY SCHEDULE, CLARK REVISION

a Ave *T*-score Menstr	*b* Ave *T*-score Premenstr	*c* Ave *T*-score Postmen	*d* Ave *T*-score Intermen	Henmon- Nelson Form *A*	Henmon- Nelson Form *B*	Thustone Personal Schedule
	57± 030	75± 030	79± 026	—21± 066	—18± 066	—.01± 068
		73± 032	71± 034	—.09± 068	—09± 068	—05± 069
			72± 033	—12± 068	—14± 067	.06± 068
				—15± 067	—19± 066	—12± 068
					80± 025	—06± 068
						—15± 067

subjects made a large number of errors, the means and the standard deviations (consequently, the T-scores) were greatly affected There were no statistically reliable differences between the average T-scores on the learning test for any one phase of the menstrual cycle and any other one of the phases of the physiological cycle, either for equations correctly completed or for equations incorrectly completed

There were no statistically reliable differences between the Forms A and B of the Henmon-Nelson Test of Mental Ability and the average T-scores on the learning tests during each of the four phases of the menstrual cycle The critical ratios between Form A of the Henmon-Nelson Test given between menstrual periods, and Form B given during menstruation, were not significant statistically Likewise there were no statistically reliable differences between learning in the four phases of the menstrual cycle and neurotic tendencies as measured by the Thurstone Personality Schedule, Clark Revision. Neurotic tendencies as indicated by this test appear to bear only a slight negative or a chance relationship to the learning test used in this experiment

Analysis was also made of the scores on the following specific days of the menstrual cycle· ($A1$) the first day of menstruation, ($A2$) the second day of menstruation, ($B4$) the last day of the premenstrual phase, ($C1$) the first day of the postmenstrual phase, ($D2$) the day midway between menstrual periods Correlations were computed between the foregoing variables, Forms A and B of the Henmon-Nelson Test of Mental Ability, and the Thurstone Personality Schedule, Clark Revision, for equations correctly completed, as given in Table 3, and for equations incorrectly completed, as given in Table 4

From Table 3 it is apparent that the correlations for the T-scores of equations correctly completed on the specific days of the menstrual cycle were high, ranging from 84± 022 to 93± 011, but varied more widely and were not as high as for the average T-scores of each of the phases (Table 1) The coefficients of correlation between the T-scores on the specific days of the menstrual cycle and the two forms of the Henmon-Nelson Test of Mental Ability were either slightly negative or approximately zero, which was also true for the Thurstone Personality Schedule There were no statistically reliable differences between intelligence and neurotic tendencies (as measured by these tests) and the T-scores of equations correctly or incorrectly completed on specific days of the menstrual cycle.

TABLE 3

CORRELATIONS OF T-SCORES ON EQUATIONS CORRECTLY COMPLETED (RIGHT) DURING FIRST AND SECOND DAYS OF MENSTRUATION (*A*1, *A*2), LAST DAY OF PREMENSTRUAL PHASE (*B*4), FIRST DAY OF POSTMENSTRUAL PHASE (*C*1), MID-DAY OF INTERMENSTRUAL PHASE (*D*2), FORMS *A* AND *B* OF HENMON-NELSON TEST, THURSTONE PERSONALITY SCHEDULE

*A*1	*A*2	*B*4	*C*1	*D*2	Henmon-Nelson Form *A*	Henmon-Nelson Form *B*	Thustone Personal Schedule
	88± 018	84± 023	95± 011	88± 016	38± 062	40± 061	— 11± 071
		87± 020	89± 017	87± 016	43± 053	17± 062	— 13± 063
			90± 015	85± 018	33± 056	38± 053	— 10± 061
				86± 017	42± 055	38± 055	— 04± 066
					38± 053	44± 050	— 13± 061
						80± 025	— 06± 068
							— 15± 067

TABLE 4

Correlations of T-Scores for Equations Incorrectly Completed (Wrong) During First and Second Days of Menstruation (*a*1, *a*2), Last Day of Premenstrual Phase (*b*4), First Day of Postmenstrual Phase (*c*1), Mid-Day of Intermenstrual Phase (*d*2), Forms *AA* and *B* of Henmon-Nelson Test, Thurstone Personality Schedule

a1	*a*2	*b*4	*c*1	*d*2	Henmon-Nelson Form *A*	Henmon-Nelson Form *B*	Thustone Personal Schedule
	49± 061	57± 051	69± 048	47± 055	05±.071	05± 071	—13± 067
		76± 034	76± 036	74± 029	—05± 064	—02± 065	—12± 063
			41± 065	58± 041	02±.062	—01± 062	—01± 062
				51± 049	—08± 066	—02± 065	.01± 065
					—18±.060	—14± 060	—14± 060
						80± 025	—06± 068
							—15± 067

As given in Table 4, the correlations for the T-scores of equations incorrectly completed on specific days of the menstrual cycle ranged from .41±.064 to .76±.036. The coefficients of correlation for equations incorrectly completed on various specific days of the menstrual cycle varied much more widely than for equations correctly completed (Table 3). The means and standard deviations of the T-scores for the specific days of the menstrual cycle and for the average T-scores of the four phases of the menstrual cycle are given in Table 5.

TABLE 5

Means and Standard Deviations of T-Scores, Correct and Incorrect Equations, on First and Second Days of Menstruation ($A1$, $A2$), Last Day of Premenstrual Phase ($B4$), First Day of Postmenstrual Phase ($C1$), Mid-Day Intermenstrual Phase ($D2$), and Average T-Score in Each of the Four Phases

		Correct equations		Incorrect equations	
	N	Mean	Stand. Dev.	Mean	Stand. Dev.
Ave. A	96	49.88	9.59	48.95	6.62
Ave. B	96	49.52	9.96	50.17	7.92
Ave. C	96	50.14	9.46	42.23	7.48
Ave. D	96	49.07	9.38	50.32	8.44
$A1$	88	49.67	10.95	48.93	10.67
$A2$	109	49.12	10.46	45.73	8.90
$B4$	117	48.75	9.79	50.33	9.31
$C1$	107	50.39	9.09	49.06	9.49
$D2$	121	49.15	10.71	49.31	11.31

Reliability of differences between means

Incorrect equations	P.E.$_{\text{diff}}$	Difference of means	Critical ratio
$B4$ and $A2$	.82	4.60	5.63
$A1$ and $A2$	.96	3.20	3.34
$D2$ and $A2$	1.03	3.58	3.47

The critical ratio between the means of the T-scores on equations incorrectly completed on the day preceding menstruation ($b4$) and the second day of menstruation ($a2$) is 5.63 and is, therefore, a statistically reliable difference. However, this is the only statistically reliable difference between the means. In Table 6 is summarized the probable errors of the differences of the means and of the coefficients of correlation, together with the critical ratios. Table 6 is based on the data given in Tables 3 and 4.

TABLE 6

THE PROBABLE ERRORS OF DIFFERENCES

(Based on Tables 3 and 4)

Probable error of difference		Critical ratio	Probable error of difference		Critical ratio
P E$_{\text{Mb4 Ma2}}$	= 817	5 6	P E$_{\text{ra1 c1 ra1 d2}}$	= 073	3 1
P E$_{\text{Ma1 Ma2}}$	= 958	3 5	P E$_{\text{ra2 b4 ra2 a1}}$	= 070	3 8
P E$_{\text{Md2 Ma2}}$	= 1 032	3 5	P E$_{\text{ra2 c1-ra2 a1}}$	= .071	3 8
P E$_{\text{rA1 A2-rA1 B1}}$	= 029	1 4	P E$_{\text{ra2 d2 ra2 a1}}$	= 068	3 7
P E$_{\text{rA1 C1-rA1 A2}}$	= 021	2 3	P E$_{\text{rb1 a2-rb4 a1}}$	= 061	3 0
P E$_{\text{rA1 C1 rA1 B1}}$	= 025	3 5	P E$_{\text{ra2 b1 rb1 c1}}$	= 073	4 8
P E$_{\text{rA1 D2 rA1 B1}}$	= 028	1 5	P.E$_{\text{rb1 a2 rb1 d2}}$	= 053	3 3
P E$_{\text{rA1 C1-rA1 D2}}$	= 020	2 4	P E$_{\text{rb1 a1 rb1 c1}}$	= 082	2 0
P E$_{\text{rA2 C1-rA2 B1}}$	= 026	8	P E$_{\text{rc1 a2 rc1 b1}}$	= 075	4 8
P.E$_{\text{rB4 C1 rB1 A1}}$	= 027	2 1	P E$_{\text{rc1 a2 rc1 a1}}$	= 060	1 2
P E$_{\text{rA1 C1 rC1 D2}}$	= 020	3 3	P E$_{\text{rc1 a1-rc1 b4}}$	= 080	3 5
P E$_{\text{rA2 IIA rA2 IIB}}$	= 082	3 2	P E$_{\text{rc1 a1 rc1 d2}}$	= 069	2 7
P.E$_{\text{rc1 a2 rc1 d2}}$	= 061	4 1	P E$_{\text{rd2 a2 rd2 a1}}$	= 063	4 3
			P E$_{\text{rd2 a2 rd2 c1}}$	= 057	4 1

From Table 6 it is evident that there are statistically reliable differences between the correlations of the *T*-scores on specific days of the menstrual cycle A critical examination of the statistically reliable critical ratios of the probable errors of the differences of the correlations of equations incorrectly completed, indicates that in the five critical ratios of 4 00 or more, and in the two critical ratios of 3 8, the constantly recurring variable is *a*2, the *T*-scores of the equations incorrectly completed on the second day of menstruation. On the second day of menstruation the destructive processes are at their height In as much as the mean of the *T*-scores on the second day of menstruation is lower than for any other of the specific days of the menstrual cycle or for the average *T*-score of any phase of the menstrual cycle (Table 5), on this day accuracy is increased

Regularity of menstruation, pain during the catamenial periods, worry and elimination have been held as factors influencing mental efficiency during functional periodicity. Therefore, a careful analysis was made of the relationships between physiological periodicity and these qualitative variables On the first day of the experiment remembered information was secured from each subject on these attributes and the contingency coefficients computed. The coefficients of contingency were so small, ranging from 04 to 30, that these attributes given from memory appear to have little significance

From each subject was also obtained a daily record of the presence or absence of emotional disturbance, pain, and worry The contingency coefficients for these attributes were computed for the specific days of the menstrual cycle and are given in Table 7

TABLE 7

CONTINGENCY COEFFICIENTS* BETWEEN EMOTIONAL DISTURBANCE, PAIN, WORRY, ON FIRST AND SECOND DAYS OF MENSTRUATION (*A*1, *A*2), LAST DAY OF PREMENSTRUAL PHASE (*B*4), FIRST DAY OF POSTMENSTRUAL PHASE (*C*1), MID-DAY OF INTERMENSTRUAL PHASE (*D*2)

Day	Emotion — Pain	Emotion — Worry	Pain — Worry
*A*1	08	65	01
*A*2	18	33	78
*B*4	85	16	.23
*C*1	03	39	42
*D*2	85	.80	15

*Corrected for broad grouping

The relationships between the attributes of pain, worry, and emotional disturbance varied widely on certain specific days in the cycle and appear to be related to physiological periodicity The high relationship between pain and worry on the second day of the catamenial period, when destruction of tissues is at its height, appears to be a factor influencing the statistically significant difference in equations incorrectly completed on the second day of menstruation

The biserial *r*'s were computed between the *T*-scores on the learning tests and presence or absence of the qualitative variables, emotional disturbance, pain, and worry, on specific days of the menstrual cycle, and are given in Table 8

From the biserial *r*'s there appears to be little relationship between the number of equations correctly completed or incorrectly completed and the qualitative variables of worry, pain, and emotional disturbance, except for a slight relationship between the presence of pain and equations incorrectly completed on the first day of menstruation

This intensive research on the effect of woman's physiological rhythm on mental efficiency as shown by the learning process, has resulted in certain significant findings The usual supposition of

TABLE 8

BISERIAL *r*'S BETWEEN CORRECT AND INCORRECT EQUATIONS AND EMOTIONAL DISTURBANCE, PAIN, WORRY, ON FIRST AND SECOND DAYS OF MENSTRUATION, LAST DAY OF PREMENSTRUAL PHASE, FIRST DAY OF POSTMENSTRUAL PHASE, MID-DAY OF INTERMENSTRUAL PHASE

	Correct equations				Incorrect equations		
Day	Emotional Dist	Pain	Worry	Day	Emotional Dist	Pain	Worry
A1	−12	17	13	*a1*	−16	−35	09
A2	04	07	16	*a2*	−12	00	09
B4	01	−14	19	*b4*	−03	22	01
C1	−05	−11	20	*c1*	27	12	07
D2	−02	−05	07	*d2*	−33	30	02

relative incompetency and inefficiency during functional periodicity is not supported by this study. Mental efficiency as indicated by the learning process is not lowered during any of the phases of the menstrual cycle. Other things being equal, functional periodicity should be no obstacle to women's success in business or in professional work which involves principally mental activities. It should be noted that this maintainance of efficiency relates to mental activity in the group as a whole, and does not refer to motor activity or efficiency in women suffering from abnormal gynaecological conditions.

In young women of this age and occupation, the physiological changes occurring during the menstrual cycle appear to have only slight or chance relationship with neurotic tendencies as measured by the Thurstone Personality Schedule, Clark Revision. The neurotic tendencies measured by this test seem to have little or no effect upon mental efficiency as shown by the learning process. The scores on mental tests taken by women subjects during the destructive stage of the menstrual cycle, will probably not be significantly affected as woman's physiological rhythm has no statistically reliable effect upon intelligence as measured by such a standardized test as the Henmon-Nelson Test of Mental Ability. The inclination to mental effort may be reduced by the physiological changes occurring during the first days of menstruation, but mental efficiency is not affected. A similar tendency has been found in studies of fatigue.

Accuracy varies during the different phases of the menstrual cycle although mental activities which require only speed are not affected by functional periodicity. The fluctuation in accuracy is outstanding on the second day of the catamenial period On this day the destructive stage of the physiological process is at its maximum, there is a high relationship between pain and worry, and there is an actual decrease in the number of errors It seems reasonable to assume that the large variation in accuracy may be either the direct result of the physiological changes or may be due to compensation for the distractions of pain and worry which, in turn, may be induced by the physiological changes. In this study, the competition of the subjects with their own past records, and with other subjects, may have supplied the motivation for increased effort resulting in compensation

Every woman should be cognizant of the fact that the physiological changes on the second day of menstruation tend to produce a depressive attitude which, in turn, tends to result in worry Thus the physiological changes condition the subjective attitudes toward objective stimuli Similarly, the physiological changes produced by certain drugs, such as opium and alcohol, by certain diseases, for example, acromegaly and diabetes, and by fatigue, likewise have a direct influence on emotional attitudes

The gynaecological changes on the second day of menstruation tend to increase accuracy, probably due to compensation, and to produce the subjective attitude of worry. To accomplish the same results, more positive motivation is needed than on other days. On this day, therefore, even though mental efficiency is not affected, women should plan their work in such a way that there will be a minimum of irritation They should recognize the fact that, if irritations do arise and there is a tendency to worry about them, the exaggeration is due to the physiological changes in progress.

CONCLUSIONS

The conclusions from this study may be summarized, as follows:

1. The validity of the learning tests was shown by the parabolic curves of learning for each subject

2 The small range in age and the number of subjects participating would indicate that the group constituted an adequate random sampling of women students between the ages of seventeen and twenty-four years, enrolled in the first three years of college

3 Learning progresses during the four phases of the menstrual cycle with no demonstrable differences on the basis of the *T*-scores

4 The correlations and intercorrelations between the average T-scores of equations correctly completed during each of the four phases of the menstrual cycle are exceedingly high.

5 The correlations and intercorrelations between the average number of equations incorrectly completed during each of the four phases of the menstrual cycle are high, but not as high as for the equations correctly completed, due to the small number of equations incorrectly completed

6. Functional periodicity has no demonstrable effect upon intelligence as measured by the Henmon-Nelson Tests of Mental Ability.

7 There is only a slight or chance relationship between learning during the four phases of the menstrual cycle and neurotic tendencies as measured by the Thurstone Personality Schedule, Clark Revision

8. The low coefficients of correlation and the high probable errors between the scores on the Thurstone Personality Schedule, Clark Revision, and on the Henmon-Nelson Tests of Mental Ability indicate that there is only a slight negative or chance relationship between scores on these tests.

9. On the basis of the T-scores, functional periodicity apparently has no effect on correct responses on specific days of the menstrual cycle, except perhaps on the first day of menstruation

10 Mental activities which require accuracy vary with functional periodicity, especially on the second day of menstruation when the destructive processes are at their height. The accuracy on the second day of menstruation is statistically reliably increased in subjects of this age and occupation.

11 Mental activities which require only speed are not affected by the gynaecological periods

12. There is only slight or chance relationship between certain mental and physical habits—regularity of menstruation, pain during catamenial periods, worry, and elimination—as given from memory.

13. Although, as given from memory, there is only a slight relationship between pain during physiological periodicity and worry, the relationship between pain and worry reported daily varies widely on specific days of the menstrual cycle The high relationship between pain and worry on the second day of the catamenial period appears to be a factor influencing accuracy

14. There is only a slight relationship between the number of

equations correctly or incorrectly completed and the qualitative attributes of pain, worry, and emotional disturbance.

15 Although the inclination to learn may be affected by the catamenial periods, mental efficiency, as measured by correct responses on the five-minute learning tests used in this study, is not affected in subjects of this age and occupation.

Physicians, sociologists, psychologists, employers in offices and factories, and the many others who find these physiological changes an extraneous variable in their dealings with women, may gain from this study a better understanding of the effect of woman's physiological rhythm upon her mental activity and efficiency Women themselves may have better insight into their own attitudes and moods during the gynaecological periods This study should aid women in making better adjustments and thus contribute to their happiness and well-being.

REFERENCES

1 ARYAMOV, I. A The influence of menstruation on behavior of girls in secondary schools *Pedologiya*, 1930, **1**, 26-30

2 CONKLIN, E S, BYROM, M. L, & KNIPS, A Some mental effects of menstruation *Ped Sem*, 1927, **34**, 357-367

3 CHADWICK, M Woman's periodicity London Doulgas, 1933 Pp 227

4 DALY, C D Der Menstruationskomplex Einer Psychoanalytische Studie *Imago*, 1938, **14**, 11-75.

5 EAGLESON, H E Periodic changes in blood pressure, muscular coordination and mental efficiency in women *Comp Psychol Monog*, 1927, **4**, No 20 Pp 65

6 ELLIS, H Man and woman London A C Black, 1926, 330-353

7 GOLDSCHMIDT, T Menstrual taboo and women's psychology *J Abn Psychol*, 1934, **29**, 218-221.

8 HALL, G S Adolescence New York Appleton, 1907, 472-512

9 HAUPTMAN, A Menstruation und Psyche (Versuch einer "verstandlichen" Ibeziehungsetzung somatischer und psychischer Erscheinungsreihen) *Arch f Psychiat*, 1924, **71**, 1-54

10. HIRSCHMANN-WERTHEIMER, I. Wechselseitige Beiziehungen von Menstruation und Psyche *Monatschr. f Psychiat u Neur*, 1927, **66**, 215-254

11 HOLLINGWORTH, L S Functional periodicity New York Teachers College, Columbia University, 1914 Pp 101

12 JOHNSON, G B The effects of periodicity on learning to walk a tight-wire *J Comp Psychol*, 1932, **13**, 133-141

13 KIRIHARA, H Functional periodicity Rep Sci of Labour Japan Kurasiki, 1932 No 14 Pp 53

14. Lewin, K., & Freund, A. Untersuchungen zur Handlungs- und Affektpsychologie. VII. Psychische Sattigung im Menstruum und Intermenstruum. *Psychol. Forsch.*, 1930, **13**, 198-217.

15. Middleton, W. C. Is there a relationship between kleptomania and female periodicity in neurotic individuals? *Psychol. Clin.*, 1933, **22**, 232-247.

16. Mosher, C. D. Functional periodicity in women and some of the modifying factors. *Collected Reprints, Johns Hopkins Univ.*, 1911, **26**. Pp. 23.

17. ———. Health and the woman movement. New York: The Woman's Press, 1918. Pp. 45.

18. ———. Woman's physical freedom. New York: The Woman's Press, 1923. Pp. 87.

19. Peaks, A. Periodic variations in efficiency. *Educ. Psychol. Monog.*, New York University, 1921, No. 23. Pp. 95.

20. Seward, J. P. The effect of repetition on reactions to electric shock; with special reference to the menstrual cycle. *Arch. of Psychol.*, 1934, No. 168. Pp. 90.

21. Seward, G. H. The female sex rhythm. *Psychol. Bull.*, 1934, **31**, 153-192.

22. Sowton, S. C. M., Myers, C. S., & Bedale, E. M. Two contributions to the experimental study of the menstrual cycle. I. Its influence on mental and muscular efficiency. II. Its relation to general functional efficiency. *Rep. Indus. Fatigue Res. Board, London*, 1928, No. 45. Pp. 72.

23. Tuttle, W. W. Changes in irritability in women during the menstrual cycle. *J. Lab. & Clin. Med.*, 1925, **11**, 60-62.

24. Voitsechovsky, N. V. (The influence of menstruation upon the nervous and psychic apparatus of women.) Thesis, Imperial Military Acad. Med., St. Petersburg, 1909, No. 6. Pp. 165.

The Mills School for Kindergarten-Primary Teachers
66 Fifth Avenue
New York City

CORRELATIONS OF CHRONOLOGICAL AGE, MENTAL AGE, AND INTELLIGENCE QUOTIENT WITH OTHER ABILITIES AND TRAITS IN GRADE I*[1]

Hunter College and Columbia University

FRANK T WILSON AND CECILE WHITE FLEMMING

INTRODUCTION

During the school year 1933-34 a variety of tests was given to 25 children in Grade I of the Horace Mann School, Teachers College These included tests of "reading readiness"; many of the Gates Reading Diagnosis tests, some reading achievement tests; mental ability tests, such as the Stanford Revision of the Binet-Simon tests and various performance tests, certain psychological tests, as of perception and perseveration, and several measures of psycho-physical and personality traits and of home background The purpose of the study was to examine any possible relationships that might exist between measurable traits and abilities, and early progress in the mechanics of reading.

The children of the group came from well-to-do homes. A large percentage of the parents were professional people The following averages for these pupils were found:

Chronological Age	6 31
Mental Age	7 61
Intelligence Quotient	120 6

Nearly every test and measurement was given or made individually, under carefully controlled conditions, and by reliable persons accustomed to administering tests to young children The co

---

*Received in the Editorial Office on November 27, 1936

[1] This report presents a minor phase of a study of Reading Readiness and Reading Progress in the Primary Grades of the Horace Mann School, Teachers College, New York, 1933-36 This study has been made possible by the cooperation of Miss Agnes Burke, Teacher of Grade I and other teachers of the Kindergarten and Primary Grades It has been made under the supervision of Doctor Cecile White Flemming, Director of Pupil Individual Development and Guidance, and of Doctor Rollo G Reynolds, Principal Prepared with the assistance of the U S Works Progress Administration, New York City, project number 65-97-295, sub-project 25

operation of the pupils was almost invariably excellent It is believed for these reasons, that errors of examination were unusually low

This report presents data in terms of correlations of chronological age, mental age, and intelligence quotient with 106 other measures and appraisals of the pupils The correlations were obtained by the rank order method To secure the rank orders all measures and appraisals were reduced to numerical scores. Owing to lack of facilities it was not feasible to make all the computations that were possible in the original study A "finder" device was used to select for computations the correlations which seemed to promise significance It is believed that through the use of this device, although it was not altogether accurate, all the high and fairly high correlations were found The correlations omitted were probably below .50, and most of them probably nearer zero than 50 The *PE* of rho's, when $N=25$, range from ± 5237 for 90 to ± 1335 for 10.

The validity of many of the measures and appraisals is uncertain Few correlations of seemingly unusual size were obtained, however, and few which were inconsistent with other correlations for the same kind of traits and abilities as found in the complete data of the original study. The opinion of the teacher, of the school psychologist, and of other qualified persons who have studied the figures, is that the results have quite high validity.

FINDINGS

Lists 1, 2, and 3 present the obtained correlations of chronological age, mental age, and intelligence quotient with the other measures in order of size As would be supposed, mental age gave a greater number of high correlations with other measures than intelligence quotient or chronological age Each of the measures gave several negative correlations. Mental age and intelligence quotient correlated 81, chronological age and intelligence quotient correlated

LIST 1—CORRELATIONS OF CHRONOLOGICAL AGE WITH OTHER MEASURES*

Gates Primary Reading Tests, Type 2, Sentence reading, March	42
Teacher's ranking in reading, November prediction	41
Gates Primary Reading Tests, Type 1, Word recognition, November	40
Healy Picture Completion Test, II	40
Hildreth First Grade Reading Analysis Test, Matching phrases	39
Metropolitan Readiness Tests, Vocabulary	37
Gates Reading Diagnosis Tests, XIII, 1-2, Write words Fall	37
Grip	36

LIST 1—CORRELATIONS OF CHRONOLOGICAL AGE WITH OTHER MEASURES (continued)*

Teacher's ranking in reading, May ability	.35
Gates Reading Diagnosis Tests, XI, 1, Repeat nonsense syllables Fall	.34
Gates Reading Diagnosis Tests, X, 2, Recognition letter sounds Fall	31
Van Wagenan Reading Readiness Test, Information	29
Metropolitan Readiness Tests, Numbers	29
Metropolitan Readiness Tests, Total	29
Gates Reading Diagnosis Tests, IX, 1-7, Phonic combinations Fall	28
Gates Reading Diagnosis Tests, XV, 2, Memory span letters Fall	28
Gates Primary Reading Tests, Type 3, Paragraph reading, March	28
Mental Age	28
Gates Primary Reading Tests, Type 2, Sentence reading, May	28
Hildreth First Grade Reading Analysis Test, Matching sentences	27
Personal Traits	06
Height	09
Gates Primary Reading Tests, Type 1, Word Recognition, March	24
Manikin	24
Hildreth First Grade Reading Analysis Test, Total	24
Gates Reading Diagnosis Tests, IX, 1-7 and 9, Giving phonic combination and letter sounds Fall	22
Gates Reading Diagnosis Tests, X, Total auditory perception Fall	22
Gates Primary Reading Tests, Type 1, Word recognition, May	22
Hildreth First Grade Reading Analysis Test, Matching words	20
Total Vocabulary time	20
Gates Reading Diagnosis Tests, X, 3-4, Giving initial and final sounds Fall	19
Gates Primary Reading Tests, Type 3, Paragraph reading, May	18
Hildreth First Grade Reading Analysis Tests, Matching words and phrases in sentences	.17
Total omissions	.17
Van Wagenan Reading Readiness Test, Word Discrimination	16
Metropolitan Readiness Test, Copying	16
Ship Test	15
Gates Reading Diagnosis Tests, XV, 4, Memory span, familiar words Fall	14
Van Wagenan Reading Readiness Test, Word learning	—.12
Tapping, score	12
Gates Reading Diagnosis Tests, XIII, 3, Write letters Fall	10
Gates Reading Diagnosis Tests, IX, 10, Recognition capital letters Fall	10
Total errors	.10
Gates Reading Diagnosis Tests, VIII, 3, Recognition words heard Fall	08
Perception	08
Perseveration, Elkins-Maller Attention Test	—08
Gates Reading Diagnosis Tests, IX, 9, Giving letter sounds Fall	07
Gates Reading Diagnosis Tests, X, 1, Blend words Fall	07
Gates Reading Diagnosis Tests, XI, 2, Same-different words Fall	07
Reversals, Visual perception	.07
Stone & Grover, Classification test for Beginners in Reading, II	.07
Total Vocabulary	05

List 1—Correlations of Chronological Age with Other Measures (continued)*

Weight	.05
Metropolitan Readiness Tests, Information	.04
Stone & Grover, Classification test for Beginners in Reading, I	.03
Steadiness	01
Van Wagenan Reading Readiness Test, Memory span	0
Metropolitan Readiness Tests, Similarities	00
Gates Reading Diagnosis Tests, XV, 1, Memory span, digits, Fall	00
Metropolitan Readiness Tests, Drawing man	—01
Gates Reading Diagnosis Tests, Recognition small letters Fall	—01
Personality Rating, Hicks, A Personality Rating Scale for Children Six to Nine	—01
Van Wagenan Reading Readiness Test, Relations	—02
Gates Reading Diagnosis Tests, XV, Total memory span Fall	—02
Stone & Grover Classification Test for Beginners in Reading, Total	—04
Metropolitan Readiness Tests, Sentences	—05
Gates Reading Diagnosis Tests, VIII, 2, Recognition word seen Fall	—05
Gates Reading Diagnosis Tests, XI, 2, and supplement, Recognition same different words Fall	—05
Mare & Foal, Time	—07
Motor co-ordination	—07
Perseveration, Elkins-Maller, Attention Test	—08
Reversals, Auditory perception	—10
Gates Reading Diagnosis Tests, XV, 3, Memory span, nonsense syllables Fall	—12
Van Wagenan Reading Readiness Test, Vocabulary	—.14
Nutrition	—18
Mare & Foal, Errors	—24
Intelligence quotient	—30
Activities, total	—48

*Some of the correlations of C A with the other measures were not computed because of limited clerical assistance. See Introduction for discussion of "finder" device.

List 2—Correlations of Mental Age with Other Measures

Intelligence Quotient	81
Metropolitan Readiness Tests, Numbers	80
Metropolitan Readiness Tests, Total	79
Total Errors	78
Gates Reading Diagnosis Tests, IX, 9, Letter sounds, Spring	75
Gates Reading Diagnosis Tests, IX, 1-7 and 9, Phonic combinations and letter sounds, Spring	73
Teacher's ranking in reading, November prediction	70
Errors, Fall	68
Gates Primary Reading Tests, Type 3, Paragraph reading, March	66
Gates Reading Diagnosis Tests, XIII, 1-2, Write words, Fall	65
Gates Reading Diagnosis Tests, VIII, 2, Word recognition, visual perception, Fall	64
Van Wagenan Reading Readiness Test, Vocabulary	64

LIST 2—CORRELATIONS OF MENTAL AGE WITH OTHER MEASURES (continued)

Van Wagenan Reading Readiness Test, Relations	62
Gates Reading Diagnosis Tests, X, 3-4, Giving initial and final sounds, Fall	61
Gates Primary Reading Tests, Type 1, Word recognition, May	61
Gates Reading Diagnosis Tests, X, 2, Recognition letter sounds, Spring	59
Gates Reading Diagnosis Tests, XV, 2, Memory span, letters, Fall	59
Hildreth First Grade Reading Analysis Test, Matching words and phrases in sentences	58
Teachers' Ranking in reading, May ability	56
Gates Reading Diagnosis Tests, Total X, Auditory perception, Fall	56
Gates Reading Diagnosis Tests, VIII, 3, Word recognition, auditory perception, Spring	56
Errors, Spring	55
Gates Reading Diagnosis Tests, XV, 1, Memory span, digits, Fall	55
Gates Primary Reading Tests, Type 2, Sentence reading, May	54
Gates Primary Reading Tests, Type 3, Paragraph reading, May	54
Total matching	54
Hildreth First Grade Analysis Tests, Matching words	54
Metropolitan Readiness Tests, Vocabulary	54
Van Wagenan Reading Readiness Test, Memory span	.54
Gates Reading Diagnosis Tests, XIII, 3, Writing letters, Spring	54
Gates Reading Diagnosis Tests, IX, 1-7, Phonic combinations, Spring	53
Gates Reading Diagnosis Tests, XV, 1-4, Total, Memory span, Fall	53
Gates Primary Reading Tests, Type 2, Sentence reading, March	53
Errors and Omissions, Total	51
Metropolitan Readiness Tests, Information	51
Gates Reading Diagnosis Tests, XI, 1, Repeat nonsense syllables, Fall	51
Gates Primary Reading Tests, Type 1, Word recognition, November	.51
Metropolitan Readiness Tests, Copying	.50
Gates Reading Diagnosis Tests, XIII, 3, Write letters, Fall	49
Gates Reading Diagnosis Tests, IX, 11, Recognition small letters, Fall	49
Gates Reading Diagnosis Tests, IX, 9, Letter sounds, Fall	49
Gates Reading Diagnosis Tests, IX, 1-7 and 9, Phonic combinations and letter sounds, Fall	49
Gates Reading Diagnosis Tests, IX, 1-7, Phonic combinations, Fall	49
Van Wagenan Reading Readiness Test, Information	48
Gates Primary Reading Tests, Type 1, Word recognition, March	48
Hildreth First Grade Reading Analysis Test, Total	48
Total letters	48
Gates Reading Diagnosis Tests, IX, 10, Recognition capital letters, Fall	47
Mare & Foal, time	47
Gates Reading Diagnosis Tests, XIII, 1-2, Write words, Spring	46
Ship Test	45
Stone & Grover, Classification Test for Beginners in Reading I	44
Vocabularies, total	41
Gates Reading Diagnosis Tests, VIII, 3, Word recognition, auditory, Fall	.40
Metropolitan Readiness Tests, Similarities	39
Total time	39

LIST 2—CORRELATIONS OF MENTAL AGE WITH OTHER MEASURES (continued)

Gates Reading Diagnosis Tests, XV, 4, Memory span, familiar words, Fall	39
Gates Reading Diagnosis Tests, VIII, 2, Word recognition, visual, Spring	38
Gates Reading Diagnosis Tests, XV, 3, Memory span, nonsense syllables, Fall	38
Gates Reading Diagnosis Tests, X, 2, Recognition sounded letters, Fall	37
Omissions, Spring	36
Gates Reading Diagnosis Tests, X, 1, Blend sounds, Fall	33
Steadiness	33
Gates Reading Diagnosis Tests, XI, 2, Same-different words, Fall	32
Healy Picture Completion Test, II	32
Seguin Form Board	31
Metropolitan Readiness Test, Drawing man	32
Hildreth First Grade Reading Analysis Test, Matching sentences	.31
Gates Reading Diagnosis Tests, XI, 2, and supplement, Total same-different words, Fall	30
Total vocabulary time	.24
Chronological Age	28
Gates Reading Diagnosis Tests, XIII, 3, Adapted, Writing Time, capitals, small letters and digits, Spring	28
Metropolitan Readiness Tests, Sentences	.29
Omissions, Total	24
Number magazines and newspapers in home	23
Omissions, Fall	.23
Gates Reading Diagnosis Tests, X, 1, Blend sounds, Spring	22
Stone & Grover, Classification Test for Beginners in Reading, Total	22
Manikin, Time	21
Mare & Foal, Errors	21
Van Wagenan Reading Readiness Test, Word learning	21
Number of children's books in home	16
Total Reversals	16
Van Wagenan Reading Readiness Test, Word discrimination	15
Reversals, Auditory perception	12
Hildreth First Grade Reading Analysis Test, Matching phrases	11
Reversals, Visual perception	08
Tapping	08
Developmental Index (babyhood)	.08
Perception	02
Stone & Grover Classification Test for Beginners in Reading, II	02
Personal Traits	02
Personality Rating, Hicks, A Personality Rating Scale for Children Six to Nine	01
Activities, Indoor	01
Perseveration, Elkins-Maller Attention Test	—02
Total Activities	—02
Grip	—05
Height	—03
Activities, Outdoor	—05
Nutrition	—12

List 2—Correlations of Mental Age with Other Measures (continued)

Motor Coordination	—14
Undesirable Traits	—16
Weight	—25
Perseveration, Elkins-Maller Attention Test	—25
Nervousness Index	—26

List 3—Correlations of Intelligence Quotient with Other Measures

Van Wagenan Reading Readiness Test, Vocabulary	74
Total Errors	68
Gates Reading Diagnosis Tests, VIII, 2, Word recognition visual, Fall	66
Errors, Fall	63
Van Wagenan Reading Readiness Tests, Relations	63
Gates Reading Diagnosis Tests, IX, 9, Letter sounds	61
Metropolitan Readiness Test, Total	59
Metropolitan Readiness Test, Numbers	59
Gates Reading Diagnosis Tests, XV, 1, Memory span, digits, Fall	57
Van Wagenan Reading Readiness Tests, Memory span	54
Mare and Foal, Time	54
Gates Reading Diagnosis Tests, X, 3-4, Initial and final sounds, Fall	53
Gates Reading Diagnosis Tests, IX, 1-7 and 9, Phonic combinations and letter sounds, Spring	53
Gates Reading Diagnosis Tests, XV, 1-4, Total memory span, Fall	53
Metropolitan Readiness Test, Information	51
Gates Reading Diagnosis Tests, IX, 11, Recognition small letters, Fall	50
Gates Primary Reading Tests, Type 3, Paragraph reading, March	49
Gates Reading Diagnosis Tests, VIII, 3, Word recognition, visual, Spring	48
Gates Primary Reading Tests, Type 1, Recognition, May	48
Total Matching	47
Gates Reading Diagnosis Tests, IX, 9, Letter sounds, Fall	47
Gates Reading Diagnosis Tests, XV, 2, Memory span, Letters, Fall	47
Hildreth First Grade Reading Analysis Test, Word and Phrases in Sentences	47
Gates Reading Diagnosis Tests, Total X, Total auditory perception, Fall	46
Gates Reading Diagnosis Tests, X, 2, Recognizing letter sounds, Spring	45
Total Letters	45
Mare and Foal, Errors	44
Gates Reading Diagnosis Tests, XV, 3, Memory span, nonsense syllables, Fall	44
Total errors and omissions	44
Gates Reading Diagnosis Tests, XIII, 3, Adapted, Writing letters and digits, Spring	44
Gates Primary Reading Test, Type 3, Paragraph reading, May	43
Errors, Spring	43
Teacher's ranking in reading, November prediction	41
Gates Primary Reading Tests, Type 2, Sentence reading, May	40

LIST 3—CORRELATIONS OF INTELLIGENCE QUOTIENT WITH OTHER MEASURES (continued)

Gates Reading Diagnosis Tests, IX, 10, Recognition capital letters, Fall	40
Metropolitan Readiness Test, Copying	40
Hildreth First Grade Reading Analysis Test, Matching words	39
Vocabulary, Total	39
Stone & Grover, Classification Test for Beginners in Reading I	38
Metropolitan Readiness Test, Similarities	38
Gates Reading Diagnosis Tests, IX, Total 1-7 and 9, Phonic combinations and letter sounds, Fall	38
Gates Reading Diagnosis Tests, XIII, 1-2, Write words, Fall	38
Gates Reading Diagnosis Tests, VIII, 3, word recognition, auditory perception, Fall	.37
Slip Test	37
Gates Reading Diagnois Tests, IX, 1-7, Phonic combinations, Spring	36
Gates Primary Reading Tests, Type 1, Word recognition, March	36
Gates Reading Diagnosis Tests, XV, 4, Memory span, familiar words, Fall	36
Gates Reading Diagnosis Tests, XIII, 3, Write letters, Fall	36
Gates Reading Diagnosis Tests, IX, 1-7, Phonic combinations, Fall	36
Metropolitan Readiness Test, Vocabulary	36
Teacher's ranking in reading, May ability	34
Gates Reading Diagnosis Tests, XI, 2, and Supplement, Same-different words, Fall	34
Number Children's Books in Home	34
Gates Reading Diagnosis Tests, VIII, 2, Word recognition visual recognition, visual perception, Spring	33
Gates Reading Diagnosis Tests, XI, Blend words, Fall	.33
Gates Reading Diagnosis Tests, XI, 2 Same-different words, Fall	32
Hildreth First Grade Reading Analysis Test, Total	31
Activities, Indoor	.31
Gates Primary Reading Tests, Type 1, Word recognition	30
Gates Reading Diagnosis Tests, X, 1, Blend sounds, Spring	29
Metropolitan Readiness Test, Drawing man	29
Gates Primary Reading Test, Type 2, Sentence reading, March	27
Activities, Total	27
Metropolitan Readiness Test, Sentences	36
Gates Reading Diagnosis Tests, X, 2, Recognition sounded letters, Fall	26
Gates Reading Diagnosis Tests, XI, 1, Repeat nonsense syllables, Fall	25
Steadiness	25
Omissions, Spring	24
Gates Reading Diagnosis Tests, XIII, 1-2, Write words, Spring	24
Van Wagenan Reading Readiness Test, Information	23
Stone & Grover, Classification Test for Beginners in Reading, Total	23
Reversals, Total	23
Total Time	.23
Number of magazines and newspapers in home	23
Van Wagenan Reading Readiness Test, Word learning	18
Reversals, Auditory perception	18

List 3—Correlations of Intelligence Quotient with Other Measures (continued)

Omissions, Total	17
Activities, Outdoor	16
Writing Time	16
Hildreth First Grade Reading Analysis Test, Matching sentences	15
Developmental Index (babyhood)	13
Omissions, Fall	12
Healy Picture Completion Test, II	11
Stone & Grover, Classification Test for Beginners in Reading, II	10
Total Vocabulary Time	09
Van Wagenan Reading Readiness Tests, Word discrimination	07
Manikin, Time	04
Personality Rating, Hicks, A Personality Rating Scale for Children Six to Nine	03
Reversals, Visual perception	03
Perseveration, Elkins-Maller Attention Test	03
Seguin, Form Board	—00
Personal Traits	—00
Motor Coordination	—02
Tapping	—04
Perseveration, Elkins-Maller Attention Test	—05
Perception	—07
Nutrition	—09
Hildreth First Grade Reading Analysis Test, Matching phrases	—14
Height	—21
Nervousness Index	—28
Chronological Age	—30
Undesirable Traits	—31
Weight	—35
Grip	—38

.30, chronological age and mental age correlated 28. These figures are about what would be expected, and agree very well with reports of previous studies On the whole the sizes of all the correlations seem reasonable with the exception, perhaps, of some of those between the various measures of mental ability. The intercorrelations of the measures of mental ability were not high

Tables 1 to 5 give the correlations of the three measures with other tests and appraisals arranged by certain groupings Table 1 shows the correlations for fourteen reading measures. The averages of the correlations are for chronological age and reading, 29, for mental age and reading, 51, for intelligence quotient and reading, .33

The average for the correlations of the intelligence quotient with the reading measures is unduly lowered by one correlation, —.14, involving the Hildreth First Grade Reading Analysis Test, Match-

TABLE 1

CORRELATIONS OF CHRONOLOGICAL AGE, MENTAL AGE AND INTELLIGENCE QUOTIENT WITH 14 READING TESTS

	CA	MA	IQ
Gates Primary Reading Tests, Type 1, Word recognition, November	40	51	30
" " " " " 1, Word recognition, March	24	48	.36
" " " " " 2, Sentence reading, March	42	.53	27
" " " " " 3, Paragraph reading, March	28	66	49
Hildreth First Grade Reading Analysis Test, Matching Phrases	39	11	—14
" " " " " Test, Matching Words	20	54	39
" " " " " Test, Matching Sentences	27	31	15
" " " " " Test, Matching Words and phrases in sentences	17	.58	47
" " " " " Test, Total	24	48	31
Teacher's Ranking in reading, November prediction	41	70	41
" " " " May ability	35	56	.34
Gates Primary Reading Tests, Type 1, Word recognition, May	22	61	48
" " " " " 2, Sentence reading, May	28	54	40
" " " " " 3, Paragraph reading, May	18	54	43
Averages	289	51	332

ing Phrases. This figure seems out of place, and is explainable only in terms of chance It seems incredible that correlations of intelligence with this test would give regularly a negative though low figure This test was the first of the Hildreth Analysis Tests given to the children and was administered successively to groups of 12 and 13. It is possible that extraneous factors such as the novelty of the task, uncertainty of as to what to do, and other conditions common when children are confronted with new and very different situations, may have been involved to such an extent as to have affected the true relationships between mental abilities and success in the test The next lowest correlation for reading and intelligence quotient was + 15 The average for the correlations of *IQ* with the

TABLE 2

CORRELATIONS OF CHRONOLOGICAL AGE, MENTAL AGE AND INTELLIGENCE QUOTIENT WITH THE TESTS OF LETTER ABILITIES

		CA	MA	IQ
Stone & Grover, Classification	Test I	03	14	38
" " "	" II	07	02	10
Gates Reading Diagnosis Tests,	VIII 2, Recognition words seen, Fall	—05	64	66
" " " "	VIII 3 Recognition words heard	08	40	37
" " " "	IX, 1-7, Phonic combinations	28	49	36
" " " "	IX, 9, Giving letter sounds	07	49	47
" " " "	IX, Total 1-7 and 9	22	49	39
" " " "	X, 1 Recognition blended words	07	33	33
" " " "	X, 2 Recognition sounded letters	31	37	26
" " " "	X 3-4, Giving initial and final sounds	19	61	53
" " " "	X Total	22	56	46
" " " "	XIII, 1-2, Writing words	37	65	38
" " " "	XIII, 3, Writing letters	12	49	36
" " " "	IX, 10, Recognition capital letters	10	47	40
" " " "	XI 2, Recognition same-different words	07	32	32
" " " "	XI, 2, plus supplement	—05	30	34
" " " "	XV, 2, Memory span letters	.28	59	47
Total Letters		—	48	45
Gates Reading Diagnosis Tests,	VIII 2, Recognition words seen Spring	—	38	33
" " " "	VIII 3, Recognition words heard	—	56	48
" " " "	IX, 1-7 Phonic combinations	—	53	36
" " " "	IX, 9, Giving letter sounds	—	75	61
" " " "	IX, Total 1-7 and 9	—	73	53
" " " "	X 1, Recognition blended words	—	22	29
" " " "	X, 2, Recognition sounded letters	—	.59	45
" " " "	XIII 1-2 Writing words	—	.46	24
" " " "	XIII 3, Adapted, Writing capitals and small letters and digits	—	54	44
	Averages	14	477	398

TABLE 3

CORRELATIONS OF CHRONOLOGICAL AGE, MENTAL AGE, AND INTELLIGENCE QUOTIENT WITH 20 MENTAL TESTS

	CA	MA	IQ
Van Wagenan Reading Readiness Test, Information	29	48	23
" " " " " Relations	—02	62	63
" " " " " Vocabulary	—14	64	74
" " " " " Memory Span	00	54	54
Metropolitan Readiness Test, Vocabulary	37	54	36
" " " Sentences	—05	29	37
" " " Numbers	29	80	59
" " " Information	04	51	51
" " " Drawing Man	—01	32	29
Gates Reading Diagnosis Tests, XI, 1, Repeat nonsense syllables	34	51	25
Vocabulary	05	41	39
Seguin Form Board	24	31	00
Mare & Foal, Time	—07	47	54
" " Errors	—24	21	44
Manikin	24	21	04
Ship Test	15	45	37
Healy Picture Completion Test, II	40	32	11
Mental Age	28	—	81
Intelligence Quotient	—30	81	—
Gates Reading Diagnosis, XV, 1-4, Total Memory Span	—02	53	53
Averages	092	472	407

TABLE 4

CORRELATIONS OF CHRONOLOGICAL AGE, MENTAL AGE, AND INTELLIGENCE QUOTIENT WITH 16 PSYCHO-PHYSICAL MEASURES

	CA	MA	IQ
Total Writing Time	—	28	16
Total Vocabulary Time	20	24	09
Perception	08	02	—07
Steadiness	01	33	25
Tapping	12	08	—04
Perseveration, Elkins-Maller Attention Test	08	—02	03
Chronological Age	—	28	—30
Grip	36	—05	—38
Motor Coordination	—07	—14	—02
Weight	05	—25	—35
Height	09	—03	—21
Nutrition	—18	—12	—09
Number outdoor activities	—	—05	16
Number indoor activities	—	01	31
Total number activities	—48	—02	27
Developmental Index (babyhood)	—	08	13
Averages	024	040	—038

TABLE 5

CORRELATIONS OF CHRONOLOGICAL AGE, MENTAL AGE, AND INTELLIGENCE QUOTIENT WITH 10 PERSONALITY MEASURES

	CA	MA	IQ
Reversals, Visual perception	07	08	03
" Auditory	—10	12	18
Total Errors	10	78	68
Total Omissions	17	24	17
Number of Children's books in home	—	16	.34
Number of magazines and newspapers in home	—	23	23
Undesirable behavior traits	26	—16	—31
Personal traits	06	02	00
Personality rating, Hicks, A Personality Rating Scale for Children Six to Nine	—01	01	.03
Nervousness Index	—	—26	—28
Averages	.079	122	107

reading measures would not, however, have been very much higher than 33, even if a fairly high positive correlation for the Hildreth matching test had been found, since thirteen other correlations were combined with this one to make up the average The table shows that, for the children studied, reading ability was somewhat but not closely related to mental age, intelligence quotient, and chronological age, and that the relative degrees of relationship were in the order named

Table 2 gives the correlations of the three measures with 27 tests of letter abilities The averages are. *M A* and letter abilities, .477; *I Q* and letter abilities, 398; and *C A* and letter abilities, 14 The averages for mental age and for intelligence quotient with letter abilities are not very different from averages of the same measures with reading tests The average of the correlations of chronological age with letter tests, 14, compares with 289, the average of the correlations of chronological age with the 14 reading tests as shown in Table 1 Because of the few number of cases and the sizes of the correlations, these differences probably have little significance. Apparently the ability of the children to recognize, write, and give names, sounds, and combinations of letters, is related to mental age, intelligence quotient, and chronological age in much the same manner and degree as is reading

Table 3 shows the correlations of the three measures with 20 mental tests of various types The averages are *MA* and mental tests 472;

IQ and mental tests 402, *CA* and mental tests 092 *MA* and *IQ* indicated fairly close relationships to the other measures with considerable variability, the range of the correlations being from .21 to 80 for *MA*, and from 00 to 74 for *IQ* It would seem that either the abilities involved in the tests are in many instances quite different, or that the reliability of the measures is very low, or more probably, that both of these explanations are true The first supposition seems significant, however, since the averages of the correlations obtained by *MA* and *IQ* with the other measures are fairly high, whereas the averages obtained by *CA* are very low, a difference in relationships which seem quite probable

Tables 4 and 5 show the correlations of the three measures with a variety of psycho-physical and personality tests and estimates That these latter measures are each concerned with very different abilities and traits appears both in the descriptive names of the abilities and traits and in the variability of the correlations. The fact that the averages are so near zero, however, seems significant In only three instances were correlations high enough to indicate possible relationships of particular importance, namely: total errors with *MA* and *IQ* 78 and 68, respectively, and total number of activities with *CA*, —48. The relationships between mental measures and errors seem logical Negative relationships between *CA* and number of activities engaged in by children have been reported in studies of both preschool and school-age children

The conclusion from the figures of these two tables seems to be that psycho-physical and personality traits and abilities are quite varied, and not closely related to chronological age, mental age or intelligence quotient.

SUMMARY

Chronological age, Binet mental age and Binet intelligence quotient gave varying degrees of correlation with 106 other measures and appraisals of 25 Grade I children When the other measures and appraisals were grouped under five general headings, the following tendencies appeared: *MA* tended to give moderate correlations with reading, letter abilities, and mental tests, *IQ* tended to give correlations somewhat smaller than *MA* with the same groups of measures, *CA* tended to give still lower correlations with reading All other correlations tended to be low or practically zero, namely, *CA* with letter abilities, mental tests, psycho-physical measures, and personality measures, except *CA* with an estimate of the number of activities which correlated —.48, *MA* with psycho-physical and personality measures, except *MA* and *IQ* with errors, which correlated .78 and .68 respectively.

Hunter College
New York City

Teachers College
Columbia University
New York City

PERSONALITY RATINGS. I RELIABILITY OF TEACHERS' RATINGS*[1]

Wesleyan University

T. A. Langlie

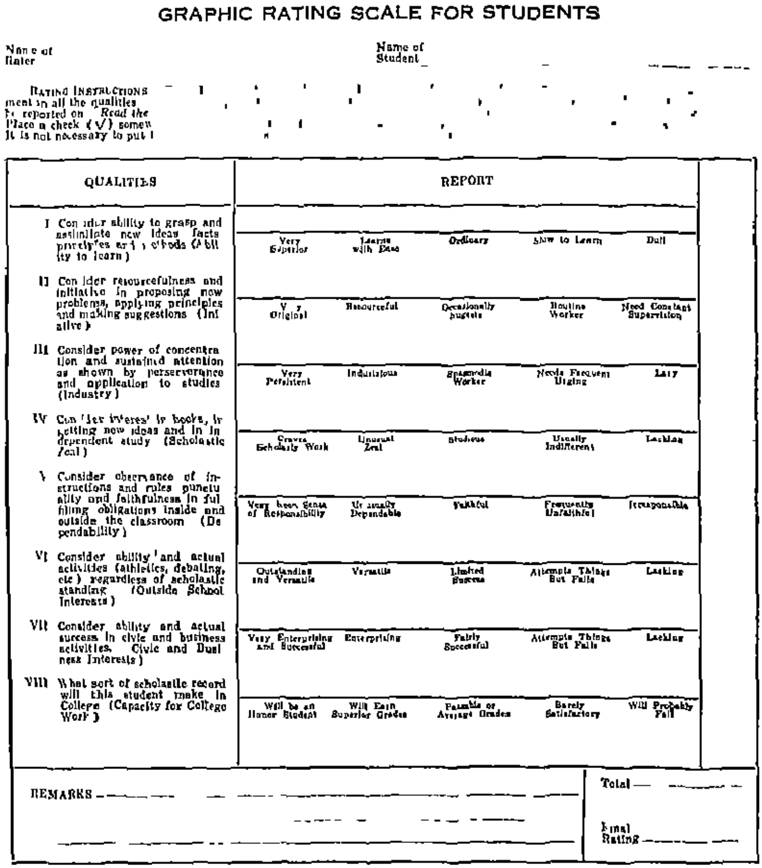
The continuing search for methods of measuring personality characteristics by the use of ratings implies a need for further careful study of the practical value of such methods in ordinary uncontrolled (non-experimental) situations. This part of a series of studies is concerned with an attempt to determine the extent to which teachers' ratings obtained under usual conditions possess the necessary reliabilities for use in predicting performance.

The Rating Scale used in this research is the usual Graphic Rating Scale, a copy of which is presented in Figure 1. The Graphic Rating Scale method has been found to be "highly reliable" when used in industrial research as reported by Paterson (4). This scale has come into general favor among personnel men and educators because of its simplicity, ease of administration, and apparent comparability of ratings. It is logically a very good method, but the reliability of ratings may be more a function of the ability of judges to observe behavior consistently and accurately than of the rating method employed. Careful and accurate judges may give reliable ratings by almost any method, and inaccurate and careless judges cannot be expected to yield reliable valid ratings by any method. Other factors remaining constant, the reliability of ratings probably will vary with methods used, and then the Graphic Rating Scale method is considered to be as accurate a method as has been devised. If results with this method are unreliable, it would appear that ratings by any method so far devised would be equally unreliable. The point is that measures of reliability of rating scales are primarily measures of the reliability

---

*Received in the Editorial Office on December 11, 1936.

[1] The first two parts of the study here reported were presented as a dissertation in candidacy for the degree of Doctor of Philosophy at the University of Minnesota in 1929. The writer is deeply indebted to Professor Donald G. Paterson for assistance in the formulation and preparation of this research.



GRAPHIC RATING SCALE FOR STUDENTS

Name of Rater

Name of Student

RATING INSTRUCTIONS
ment in all the qualities
be reported on Read the
Place a check (√) somew
It is not necessary to put

QUALITIES	REPORT				
I Consider ability to grasp and assimilate new ideas, facts, principles and methods (Ability to learn)	Very Superior	Learns with Ease	Ordinary	Slow to Learn	Dull
II Consider resourcefulness and initiative in proposing new problems, applying principles and making suggestions (Initiative)	Very Original	Resourceful	Occasionally Suggests	Routine Worker	Need Constant Supervision
III Consider power of concentration and sustained attention as shown by perseverance and application to studies (Industry)	Very Persistent	Industrious	Spasmodic Worker	Needs Frequent Urging	Lazy
IV Consider interest in books, in getting new ideas and in independent study (Scholastic Zeal)	Craves Scholarly Work	Unusual Zeal	Studious	Usually Indifferent	Lacking
V Consider observance of instructions and rules, punctuality and faithfulness in fulfilling obligations inside and outside the classroom (Dependability)	Very keen Sense of Responsibility	Unusually Dependable	Faithful	Frequently Unfaithful	Irresponsible
VI Consider ability and actual activities (athletics, debating, etc) regardless of scholastic standing (Outside School Interests)	Outstanding and Versatile	Versatile	Limited Success	Attempts Things But Fails	Lacking
VII Consider ability and actual success in civic and business activities (Civic and Business Interests)	Very Enterprising and Successful	Enterprising	Fairly Successful	Attempts Things But Fails	Lacking
VIII What sort of scholastic record will this student make in College (Capacity for College Work)	Will be an Honor Student	Will Earn Superior Grades	Passable or Average Grades	Barely Satisfactory	Will Probably Fail

REMARKS

Total

Final Rating

FIGURE 1

of the *judgments of the raters* Reliability is not as much a function of the method used as of the ability of the judge to rate

In this research rating scales were distributed to teachers in three large high schools in Minneapolis and St Paul shortly before the June commencement, with the request that they rate those students assigned to them with whom they had sufficient acquaintance No special training or instructions were given them, since *usual* conditions were desired.

The number of senior students and raters in each school are presented in Table 1

TABLE 1

High School	Year	No of Seniors	Total No of Raters	No of Teachers Who Rated 20 or More Students
Mpls , West	1924	345	48	12
St Paul, Central	1925	270	35	14
St Paul, Central	1926	290	30	13
St Paul, Humboldt	1925	90	24	0

Each scale was scored from 1 to 10 in each trait and totals were obtained for all traits excluding Traits VI and VII

In determining the reliability of the ratings collected for this study we shall discuss the following factors·

1 Normality of distributions of ratings as shown by presence or absence of skewing.
2 Spread of ratings
3. Variations in rating standards
4 Variations in rating tendencies from one year to another
5 The factor of sex in conditioning ratings.
6 The effect of "correcting" for variability of teachers' rating tendencies
7 "Halo" as a factor

1 NORMALITY OF DISTRIBUTIONS OF RATINGS AS SHOWN BY PRESENCE OR ABSENCE OF SKEWING

The assumption that the normality of distributions of ratings is a criterion of the reliability of ratings is sound provided there have been a sufficient number of subjects rated by each judge, and provided sampling errors have been avoided This criterion is probably the weakest criterion of reliability of the ratings obtained in this research, since we cannot be sure that sampling errors have been avoided and the number of subjects rated by each teacher ranges from only 20 to 101, with a mean of 41 5

"Normality" as discussed in this section refers only to lack of skewness, since "spread of ratings" will be considered in the next section

The following procedure was employed in order to obtain crude measures of skewness. The differences between Median and Third

Quartile ratings of each teacher were obtained. These will be referred to as Q_3—Mdn. Next, the differences between the Median and First Quartile ratings were obtained. These measures will be referred to as Mdn—Q_1. These two measures of variability were then compared to determine scattering above or below the Median. When Q_3—Mdn is greater than Mdn—Q_1 we know that that distribution is skewed positively, i.e., the cases are bunched below the Median. When Q_3—Mdn is smaller than Mdn—Q_1 the distribution is skewed negatively, i.e., the cases are bunched above the Median. In order to show these relationships as well as to isolate "skewness," ratios were obtained between these two measures of variability. It is apparent from this formula that a ratio of 1.00 indicates that the distribution is approximately normal, i.e., ratings are not bunched on either side of the Median, but are scattered equally on both sides of the Median. A ratio greater than 1.00 indicates that there is more scatter above the Median than below, or that the distribution is skewed positively. A ratio of less than 1.00 indicates the opposite type of distribution.

The following criterion of normality was used. If the range from the Median score to either Quartile score is not more than 20 per cent greater than the range from the Median to the other Quartile point, skewness will be considered to be negligible. That is, when Q_3—Mdn is not more than 120 per cent of Mdn—Q_1, the distribution will be considered to be approximately normal. Similarly, when Mdn—Q_1 is less than 83 per cent of Q_3—Mdn, the distribution will be considered to be approximately normal.

The frequency of various degrees of skewness is shown by the figures in Table 2. In this table the Roman Numerals refer to the Personality traits included in this study, and the ratios refer to degrees

TABLE 2

FREQUENCY OF VARIOUS DEGREES OF "SKEWNESS" FOUND IN DISTRIBUTIONS OF RATINGS OF PERSONALITY TRAITS

Nature of Skewness	$\frac{Q_3\text{—Mdn}}{\text{Mdn—}Q_1}$ Ratio × 100	Personality Traits I	II	III	IV	V	VIII
Negative	1 — 66	13	10	20	6	5	10
	67 — 82	5	2	6	3	4	3
Normal Dist	83 — 120	7	10	7	10	12	16
Positive	121 — 150	0	7	4	8	6	4
	151 — 1000	14	10	2	12	12	6

of skewness. Three degrees of skewness are shown in this table, namely normal distribution, with practically no skewness, positive and negative skewness of medium degree, and marked skewness, both positive and negative.

The above data suggest (1) that teachers seldom distribute their personality trait ratings normally over the scale, (2) that the form of distribution of ratings varies according to the traits being judged, as represented by (*a*) lack of discrimination of differences between students who are average or above in "Industriousness" (III), but relatively marked differentiation of those who are deficient in their possession of this trait, (*b*) lack of differentiation between students who possess "average" or lower degrees of "Scholastic Zeal" (IV) and "Dependability" (V), associated with relatively better differentiation among those who are superior in their possession of these traits. In other words, according to teachers' estimates of the personality traits of their pupils, most students are industrious, but some are very deficient in their possession of this characteristic The deficients are separated from the rest, but the outstandingly industrious students are merged with the average students On the other hand, most students have only a "fair" amount of "Scholastic Zeal" or are "faithful" but some students are noticeably outstanding for their "zeal" and some are "unusually dependable," while inferior students in both traits are merged with average students

In so far as normality of distributions of ratings is a valid criterion of reliability of ratings, the results reported above suggest a lack of reliability, since skewed distributions are the most prevalent type, and since there is a lack of discrimination by teachers of equal degrees of at least three of the traits included in this study

2. Spread of Ratings

Other factors remaining constant, those ratings are most valuable which are spread over the greatest range, because such ratings allow for finer discrimination among the individuals rated The Graphic Rating Scale allows for as fine a discrimination as the rater cares to make

If we assume that the individuals rated possess the various characteristics in randomly varying amounts, then we should expect a normal distribution of the trait ratings, provided we have avoided errors of sampling We should then expect that the entire scale would cover

7 Probable Errors, and that the middle 50 per cent would cover 2 Probable Errors, that is, the range between Q_3 and Q_1 would be about 2.86 units of the scale since the scale is composed of 10 units. We found above that ratings were usually skewed, and we shall now see whether there is a normal amount of scatter.

Inter-quartile ranges are summarized in Table 3.

TABLE 3

MEANS AND SD'S OF THE INTER-QUARTILE RANGES

High School		Personality Traits I	II	III	IV	V	VIII
West	Mean	2.19	3.24	2.13	1.77	1.90	2.06
	S.D.	61	87	70	.86	74	53
Central 1925	Mean	2.42	2.63	2.07	1.97	1.89	2.29
	S.D.	52	64	69	81	76	62
Central 1926	Mean	2.02	2.20	2.19	2.07	2.13	2.00
	S.D.	.35	46	77	64	98	48

From an inspection of the measures in Table 3 we find that, on the average, there is a tendency to crowd the middle 50 per cent of ratings into a range from 1.77 to 3.24 units with an average of 2.18 units instead of 2.85. This may appear to be a small difference, and yet the average spread is only 76 per cent of the theoretical spread.

As actually used, however, these scales are really composed of only seven steps, since very few ratings of 1, 2, or 10 are made. If we consider that we are really dealing with seven step scales, then the middle 50 per cent of cases should fall within two scale steps. Actually, they fall within 2.18 steps, as stated above. This result suggests that, on the average, teachers do spread their ratings to a "normal" extent. In other words, individual differences are recognized by teachers' ratings to a degree equal to that which might be expected by theory. Such are the results when dealing with average scatter.

Marked variabilities in indicating differences are found, however, when individual teachers are studied. The extent of these variabilities is shown by the standard deviations which accompany the average inter-quartile ranges in Table 3. These sigmas range from 35 to 98, and average 67. The coefficient of variability, obtained by dividing 67 by 2.18, the mean range, is 31. This figure suggests that the variability of the interquartile ranges obtained from distributions of these teachers' ratings is quite marked, and that an average figure is, accordingly, misleading. Some raters tend to cluster their estimates very closely around the mean while other teachers scatter their ratings with about equal frequency over most of the scale.

The amount of spread, furthermore, varies with the traits being rated. The possession of "Initiative" (II), for example, is considered to vary more than any other by the three groups of teachers. "Scholastic Zeal" (IV) and "Dependability" (V) are the least variable on the average, according to these ratings.

It appears that even with the Graphic Rating Scale we do not have comparable scales when used by several raters, since some teachers spread their ratings over a range seven times as great as that over which other teachers spread their ratings. Of course, we must remember that part of this variability may be in accordance with true variations in the composition of different groups of students. However, it is doubtful if this fact would account for the extreme variability here reported.

3. VARIATIONS IN RATING STANDARDS

This criterion is concerned with the comparability of the scales used by various raters to a greater degree than are the preceding criteria. We have found that "normal" distributions of ratings are rare, and that there is little agreement among raters as to the amount of "spread" of these traits. Such findings suggest that there is a tendency for teachers to differ among themselves with respect to "severe," "moderate," or "lenient" rating standards. Data permitting analysis of such tendencies are contained in Table 4, which presents the means of the Median ratings together with the Standard Deviations of their distributions.

TABLE 4

MEANS OF THE MEDIANS

High School		Personality Traits I	II	III	IV	V	VIII
West	Mean	6.65	5.90	7.21	5.66	6.06	6.09
	S.D.	.65	.55	.25	.39	.73	.46
Central	Mean	6.53	5.86	7.22	5.80	6.20	6.16
1925	S.D.	.84	1.04	.41	.58	.66	.50
Central	Mean	6.71	6.28	7.18	6.00	6.55	6.50
1926	S.D.	.70	.46	.59	.74	.62	.43

Probably the most noticeable characteristic of the means in Table 4 is the agreement among the three groups of teachers that "Industry"

(III) is possessed to a greater extent than any of the remaining traits, assuming that the scales for the six traits are absolutely equivalent On the other hand "Scholastic Zeal" (IV) is thought to be possessed to the least extent In general, there is marked agreement among the three groups of teachers as to the relative possession of these six traits In other words, it appears that high school teachers have fairly uniform standards of judgment as to the degree to which their students, on the average, are characterized by these traits.

But what of the individual judge and his standards as related to the group standard? The measures of variability (*SD* of distribution of Medians) presented in Table 4 suggest a considerable degree of inconsistency, the Standard Deviations ranging from 25 to 1 04 with an average of 59 The magnitude of this average may be emphasized by pointing out that .59 is about 60 per cent of the semi-inter-quartile range of ratings A certain variability should be expected, however, even though these ratings be perfectly reliable, because there are true differences among the groups rated To get a measure of these differences, the College Ability Test scores (presumably measuring "Ability to Learn") of most of the 1925 Central groups were obtained, and the Median score of each of 12 of these groups compared with the Median "Ability to Learn" rating (I) of each group The coefficient of correlation (Rank Order Difference Formula) which resulted was only +0 37, indicating very little relationship between Median ratings by teachers and the Median College Ability scores The differences between Median estimates therefore, do not represent true group differences. It would seem that they are due instead to lack of comparable standards

4. VARIATIONS IN RATING TENDENCIES FROM ONE YEAR TO ANOTHER

We have seen that there are marked variations in raters' standards We shall now inquire into the permanence of the rating tendencies of individual teachers. The question is,—do teachers maintain the same tendencies to rate high, medium, or low from one year to the next? We shall attempt to determine this by comparing the ratings made by the same teachers one year later. These are raters numbered 28, 29, 35, 38, 71 and 93 The ratings which will be compared are the ratings on Trait I, "Ability to Learn" The Median and Quartile ratings of these six teachers on Trait I (Ability to Learn) with the

Median and Quartile College Ability Test scores of the same groups are presented in Table 5.

TABLE 5

MEASURES OF CENTRAL TENDENCY AND VARIABILITY OF TRAIT I, AND INTELLIGENCE SCORES, 1925 AND 1926, ST PAUL CENTRAL RATINGS

Rater No	1925 N	1925 Q_3	1925 Median	1925 Q_1	1926 N	1926 Q_3	1926 Median	1926 Q_1
28	27	7.45	5.68	5.20	28	7.56	6.67	5.50
29	59	8.05	7.22	5.71	33	7.65	7.10	5.58
35	55	6.88	5.94	5.39	38	7.68	7.20	6.21
38	25	9.31	8.15	6.13	29	7.78	7.30	6.31
71	38	7.25	5.70	5.28	42	7.50	6.00	5.30
93	47	8.53	7.56	6.44	75	8.32	7.35	5.85
			Intelligence Scores					
28	27	68.50	40.29	32.79	28	83.86	71.00	44.33
29	59	83.00	71.83	46.00	33	88.50	66.00	44.13
35	55	55.00	34.64	22.00	38	66.63	48.86	37.88
38	25	88.50	71.00	48.50	29	82.67	61.00	53.50
71	38	57.07	46.83	28.92	42	70.17	46.00	29.13
93	47	49.33	39.33	22.25	75	73.81	58.22	43.50

An analysis of these figures confirms our former conclusion that these ratings are not comparable and hence not reliable.

Teacher No. 29 rated 59 students in 1925, yielding a Median rating of 7.22, a Q_3 rating of 8.05 and a Q_1 rating of 5.71. In 1926 this teacher rated 33 students, yielding a Median rating of 7.10, a Q_3 rating of 7.65 and a Q_1 rating of 5.58. This rater apparently maintained his tendency to rate fairly high from one year to the next, but as measured by the University of Minnesota College Ability Test the 1925 group rated by this teacher had a Median percentile score of 71.83, a 25th percentile score of 46.00 and a 75th percentile score of 83.00, whereas the group rated by this teacher in 1926 had a Median intelligence score of 66.00, a 25th percentile score of 44.13 and a 75th percentile score of 88.50. It appears that this rater maintained approximately the same median tendency even though the groups rated differed both in central tendency and in spread, as well as in the nature of the distribution. The estimates indicate that the 1925 group is slightly more heterogeneous than the 1926 group and that the distribution is skewed negatively for both groups. The test scores indicate that the 1926 group is the more heterogeneous and that this distribution is approximately normal. We here have evidence that although the true averages of the two groups differ, the rater may assign them approximately the same averages.

Similar analyses can be made of the other figures in Table 5, supporting the conclusion that ratings tend to remain constant even when the groups differ.

Correlations bring out this conclusion even more forcibly. The Rank Order Difference coefficients of correlation between Median ratings and Median test scores are +0.26 and +0.23 for the 1925 and 1926 groups, respectively. As obtained before with 12 raters this correlation was +0.37. When obtained by combining the ratings for both years the correlation is +0.42. The correlation coefficient between the test scores of the 1925 groups and the test scores of the 1926 groups is +0.37. The correlation between the Median ratings of the 1925 and 1926 groups is +0.83± .098. Evidently these raters tend to give the same Median ratings from one year to the next, but their rating tendencies do not vary directly with variations in the tested ability of the groups rated. That is, a group average may change, but the rated Median remains fairly constant. In other words, these results suggest that teachers are unable to differentiate differing groups when the ratings are made one year apart.

5. THE FACTOR OF SEX IN CONDITIONING RATINGS

The reliability of ratings may be a function, among other factors, of the sex of the rater as well as the sex of the subject. If teachers of different sex tend to rate all subjects differently, and if teachers of one sex tend to judge subjects of the opposite sex differently than they do subjects of the same sex, then analyses of ratings that do not take into account such relationships would be relatively unreliable. The data for our analysis of sex differences among these raters were obtained from the Central High School ratings, and are presented in Tables 6 and 7.

Table 6 presents measures of central tendency and variability of boys and girls according to whether they were rated by men teachers or women teachers. Differences between Median ratings, the Probable Errors of the Medians and of the Differences, and Coefficients of the Reliability of the Differences are also presented. For example, eight men teachers rated 84 boys and 85 girls on "Ability to Learn" (I), assigning them Median ratings of 7.06 and 7.23, respectively. The Reliability of the Medians is represented by PE_{Mdn} and the Reliability

TABLE 6

DIFFERENCES BETWEEN RATINGS OF BOYS AND GIRLS ACCORDING TO SEX OF TEACHERS CENTRAL HIGH SCHOOL, 1925

Rated By	N	Trait	*Boys* N	Median	PE_{Mdn}	Q	*Girls* N	Median	PE_{Mdn}	Q	Diff *	PE_D	$\frac{D}{PE_D}$
Men	8	I	84	7 06	15	1 09	85	7 23	19	1 37	17	24	71
Women	22	I	138	5 87	13	1 23	304	7.07	09	1 22	1 20	16	7 50
Men	8	II	84	5 53	24	1 74	85	5 72	25	1 86	19	35	54
Women	22	II	138	6 83	19	1 71	302	5 88	11	1 48	19	22	86
Men	8	III	84	7 05	16	1 21	84	7 49	06	47	44	17	2 59
Women	22	III	138	5 69	14	1 25	303	7 47	05	62	64	15	4 27
Men	8	IV	83	5 51	09	60	85	5 89	14	1 01	38	17	2 24
Women	22	IV	132	5 38	11	1 06	295	5 72	08	1 10	34	.14	2 43
Men	8	V	84	5 58	08	55	85	5 91	15	1 11	33	17	1 94
Women	22	V	137	5 55	.09	78	300	5 88	09	1 22	33	13	2 54
Men	8	VIII	84	5 80	14	1 05	85	6 84	16	1 17	1 04	21	4 95
Women	22	VIII	138	5 59	11	1 10	301	6 35	09	1 21	76	14	5 43
Men		C A T	60	53 00	3 11	19.26	75	47 79	3 49	23 53	—5 21	4 61	—1 13
Women		C A T	79	53 27	2 09	14 82	131	47 30	2 20	20 20	—5 97	3 03	—1 97

*Note Under the column headed "Diff" positive numbers indicate differences in favor of the girls and negative numbers indicate differences in favor of the boys

TABLE 7

DIFFERENCES BETWEEN RATINGS BY MEN TEACHERS AND WOMEN TEACHERS OF PUPILS CLASSIFIED BY SEX CENTRAL HIGH SCHOOL, 1925

		Rating by 8 men teachers				Rating by 22 women teachers						
		N	Median	PE_{Mdn}	Q	N	Median	PE_{Mdn}	Q	Diff*	PE_D	$\frac{D}{PE_D}$
Boys	I	84	7 06	15	1 09	138	5 87	13	1 23	—1 19	19	—6 26
	II	84	5 53	24	1 74	138	5 69	19	1 71	16	31	52
	III	84	7 05	16	1 21	138	6.83	14	1 25	— 22	21	—1 05
	IV	83	5 51	09	60	132	5 38	11	1 06	— 13	14	— 93
	V	84	5 58	08	55	137	5 55	09	78	— 03	.12	— 25
	VIII	84	5 80	14	1 05	138	5 59	11	1 10	— .21	.18	—1 17
	C A T	60	53 00	3 11	19 26	79	53 27	2 09	14 82	27	3 75	07
Girls	I	85	7 23	19	1 37	304	7 07	09	1 22	— 16	21	— 76
	II	85	5 72	25	1 86	302	5 88	11	1 48	16	27	59
	III	84	7 49	06	.47	303	7 47	05	62	— 02	08	— 25
	IV	85	5 89	14	1 01	295	5 72	08	1 10	— 17	16	—1 06
	V	85	5 91	15	1 11	300	5 88	09	1 22	— 03	17	— 18
	VIII	85	6 84	16	1 17	301	6 35	09	1 21	— 49	18	—2 72
	C A T	75	47 79	3 40	23 53	131	47 30	2 20	20 20	— 49	4 05	— 12

*Note: Under the column headed "Diff" positive numbers indicate differences in favor of the girls and negative numbers indicate differences in favor of the boys of the Difference between these Medians is represented by the PE_D, the difference (D) between the Medians is considered to be statistically significant when the ratio $\frac{D}{PE_D}$ is equal to 4 00

The rated difference between boys and girls on Trait I, when the judges were men teachers, is only 17 of a scale step The PE_D is 24 The ratio, then, is 71, which indicates that this difference is not statistically significant

According to the figures presented in this table, there are very few statistically significant differences between the ratings of boys and girls by men and women teachers Men teachers rate girls superior to boys in all of the traits, but the difference is statistically significant only in their ratings on "Capacity for College Work" (VIII) Women teachers judge girls to be superior to boys in all the traits enumerated, and these differences are statistically significant for three traits, "Ability to Learn" (I), "Industry" (III), and "Capacity for College Work" (VIII). Women teachers judge the differences between boys and girls to be larger, relative to the ratings of men teachers, in but two traits, "Ability to Learn" (I), and "Industry" (III)

The tested ability of boys and girls is represented by College Ability Test scores. According to these scores there is no significant difference between the sexes, but the tendency is in favor of the boys This is opposed to teachers' estimates on "Ability to Learn" (I)

Such differences have been reported in other investigations Cunningham (1), Hart and Orlando (2), and Paterson and Langlie (5) have all called attention to the "sex halo," which is a tendency to rate girls higher than boys, despite objective test records Paterson and Langlie, for example, pointed out that, whereas 60 per cent of the boys in high school senior classes equal or exceed the Median score of girls on "Intelligence" tests, only 40 per cent of boys equal or exceed the Median score of girls in high school scholarship They suggested that this paradoxical situation is due in part to a tendency to overestimate the performance of girls relative to boys, because of certain "personality factors" such as "Scholastic Zeal," "Dependability," etc. Furthermore, they presented evidence to show that some such factor is operative, by demonstrating that the use of "objective" achievement examinations reduces and sometimes changes the direction of differences in achievement

It is suggestive that, according to Table 6, although the differences between boys and girls in "Dependability" (V) are not statistically significant, boys tend to cluster very closely about their Median rating, while girls scatter over the scale to an extent about twice as great The *Q*'s of the distributions of boys are .55 and 78, whereas the same measures for girls are 1 11 and 1 22 The opposite tendency is found *in their possession of "Industry" (III). These facts suggest that* both men and women teachers judge boys to be uniformly "average" in "Dependability" (V), but scattered over the scale in their possession of "Industry" (III) Girls are uniformly "high" in their possession of "Industry" (III), but scattering in their possession of "Dependability" (V)

We found from the data of Table 6 that women teachers tended to rate the sexes as different in their possession of the traits studied to a greater extent than did the men teachers, but that most of the differences were not significant, and one at least was contrary to fairly well established facts The data of Table 7 enable us to determine the extent to which men and women teachers agree in the ratings of boys and again of girls For example, we find that male judges assign *boys a Median Intelligence rating of 7 06 Female judges assign a* Median rating of 5 87 or 1 19 less than that assigned by the men, despite the fact that the two groups of men are practically identical in Median tendency when measured by an intelligence test. This difference in ratings is the only statistically significant difference found, although there is a suggestion that women assign slightly lower ratings on the average whether rating boys or girls

The findings reported in this section support the conclusion that personality ratings are influenced by a factor which may be called the "sex halo" factor This is a tendency to rate girls uniformly higher in personality traits, even when opposed to objective test scores, and probably accounts in part for the paradoxical situation of a sex difference in scholarship in favor of girls, as opposed to a sex difference in tested capacity, in favor of boys Since most high school teachers are women this tendency is even more pronounced than it might be if half of the teachers were men, since, although men tend to be more lenient in rating both sexes, women are more lenient in rating girls than in rating boys

It is the opinion of the writer, however, that sex is not an extremely important factor in determining the reliability of ratings, except, possibly, for ratings in "Ability to Learn" (I), because of the marked variability of rating tendencies within both sexes. The individual differences within either sex are more pronounced than are the differences between sexes.

6. THE EFFECT OF "CORRECTING" FOR VARIABILITY OF TEACHERS' RATING TENDENCIES

If the rating by two judges on the same students are alike in every respect, except for location on the scale, we can treat the ratings statistically by the Scott Company Method (4), and obtain reliable and comparable results.

The West High School teachers included in this study rated, on the average, 47 students, a number sufficiently great to justify an attempt to determine the effect of a correction for differences in rating tendencies.

The ratings on each trait were corrected by the Scott Company Method, which may be described, as used in this research, as follows: Frequency distributions of each teacher's ratings on each trait, as well as total scores, are plotted and divided into five parts. The upper 10 per cent of ratings on each distribution are assigned to a letter grade of *A*, the next 20 per cent, a letter grade of *B*, the next 40 per cent, *C*, the next 20 per cent, *D*, and the lowest 10 per cent are given an *E* rating. In this manner, differences in tendencies to rate "high," "medium," or "low" presumably are equated. The assumption underlying this method of correction is that each group of subjects is a random collection of individuals, or at least selected in the same manner as every other group, and that they are equal in their possession of the personality traits included in the scale.

Letter scores may be used statistically by assigning values from 5 to 1 for the letters *A* to *E*, respectively. This was done in this research and correlations were then obtained between these "corrected" scores and "raw" scores. These correlations (Pearson *r*) are presented in Table 8.

When scores are corrected in the manner described above, there are very slight differences between the raw scores and the corrected scores as shown by the correlations of about .90. The total scores in particular do not change significantly when corrected. As might have been expected those ratings are least alike which are rated least reliably as measured by scattering of Median ratings (See Standard Deviations of Table 4).

TABLE 8

COEFFICIENT OF CORRELATION BETWEEN "RAW" AND "CORRECTED" SCORES (WEST HIGH SCHOOL)

Trait	Girls r	Girls PE	Boys r	Boys PE	Both r	Both PE
I	.86	010	90	009	88	.007
II	94	005	84	014	90	006
III	.94	005	92	007	93	004
IV	88	009	90	006	90	006
V	83	012	85	013	84	009
VIII	93	005	89	010	91	005
Total	94	009	94	010	.94	007

Correcting ratings statistically in the manner described does not seem to change the ratings to any marked extent since "corrected" ratings correlate very highly with uncorrected or "raw" scores In fact, the correlation between "corrected" and "raw" scores is much higher than the reliability coefficients of either kind of rating score.

This fact leads us to conclude that tendencies to rate "too high" or to rate "too low" are not sufficiently pronounced to necessitate a correction procedure. Such a conclusion is somewhat surprising in view of prevalent expectations that any group of teachers will almost inevitably contain some who are "easy markers" and some who are "hard markers" It is possible, of course, that such a tendency does not disclose itself in the initial stages of such a novel method of grading students as is afforded by a personality rating scale Perhaps, if such ratings were demanded as a routine duty, such differential rating tendencies would emerge at a later date as attitudes of harshness or sympathy would have opportunity to creep into the whole rating procedure

7 "HALO" AS A FACTOR

Other factors remaining constant, that rating method is the weakest which yields high correlations between obviously unrelated traits The data for a study of this factor include ratings from all three schools West, Central (1925), and Humboldt.

Intercorrelations between ratings on each trait were obtained to show the extent to which they represent the same characteristics These intercorrelations were obtained by correlating ratings on Trait I, for example, with ratings on Trait II for a group of pupils as afforded by a number of different teachers. Since some pupils were rated by more than one teacher such pupils' ratings are represented more than once. But we are concerned here with the judgments that have been made. The correlations presented, accordingly, show the degree of relationship between ratings by individual teachers on different traits.

We might have averaged the ratings of two or three teachers in order to increase the reliability of ratings for each trait, but the number of students who were rated by more than one teacher was not large enough to warrant such a procedure Furthermore, such a method of obtaining intercorrelations would probably increase the coefficient of correlation The measures were not corrected for attenuation, since no true reliability coefficients were available, no teacher having rated the same pupils at two different times

Bearing in mind the fact that if our measures were more reliable the intercorrelations would be higher, we will turn to Tables 9, 10 and 11, which present coefficients of correlation, showing the relationship between unaveraged estimates of each trait

TABLE 9

INTERCORRELATIONS OF RATINGS OBTAINED AT WEST HIGH SCHOOL

Traits	I	II	III	IV	V	VIII
		Boys, 66 ratings on 47 boys				
I		63	39	49	29	56
II	63		47	51	46	52
III	39	47		72	94	72
IV	49	51	72		80	71
V	29	46	94	80		71
VIII	56	52	72	71	71	
Average	47	52	65	65	64	6
		Girls, 74 ratings on 47 girls				
I		79	73	75	72	85
II	79		70	88	68	79
III	73	70		80	79	82
IV	75	88	80		81	83
V	72	68	79	81		78
VIII	85	79	82	83	78	
Average	77	77	77	81	76	81

High degrees of intercorrelation between personality ratings on different traits may mean one or both of two things (*a*) the scale used is poorly constructed since the traits being rated overlap to a marked extent, or (*b*) the raters have not analyzed their subjects sufficiently, but have judged them in each trait according to their own generalized attitudes toward them

TABLE 10

INTERCORRELATIONS OF RATINGS OBTAINED AT CENTRAL HIGH SCHOOL, 1925

Traits	I	II	III	IV	V	VI	VII	VIII
Boys, 210 ratings on 80 boys								
I		78	64	64	43	55	48	83
II	78		58	70	39	56	69	71
III	64	58		65	50	28	31	60
IV	64	70	65		48	59	55	61
V	43	39	50	48		27	42	40
VI	55	56	28	59	27		84	49
VII	48	69	31	55	42	84		63
VIII	83	71	60	61	40	49	63	
Average	62	63	51	60	41	51	56	61
Girls, 290 ratings on 134 girls								
I		72	64	67	57	55	57	80
II	72		58	78	63	64	64	69
III	64	58		70	60	39	48	65
IV	67	78	70		77	58	64	76
V	57	63	60	77		53	72	70
VI	55	64	39	58	53		77	59
VII	57	64	48	64	72	77		70
VIII	80	69	65	76	70	59	70	
Average	65	67	58	70	65	58	65	70

TABLE 11

INTERCORRELATION OF RATINGS OBTAINED AT HUMBOLDT HIGH SCHOOL

Traits	I	II	III	IV	V	VI	VII	VIII
Boys, 30 ratings on 30 boys								
I		61	55	49	54	08	45	59
II	61		81	59	65	43	70	63
III	55	81		67	73	23	64	75
IV	49	59	67		67	34	53	70
V	54	65	73	67		33	54	76
VI	08	43	23	34	33		70	17
VII	45	70	64	53	64	70		44
VIII	59	63	75	70	76	17	44	
Average	47	63	63	57	62	33	59	58
Girls, 60 ratings on 60 girls								
I		90	50	55	52	79	64	95
II	90		35	62	50	82	73	76
III	50	35		49	66	23	73	56
IV	55	62	49		59	53	53	68
V	52	50	66	59		41	51	66
VI	79	82	23	53	41		72	74
VII	64	73	73	53	51	72		63
VIII	95	76	56	68	66	74	63	
Average	69	67	50	57	55	61	64	71

The construction of the scale upon which ratings were obtained in this study argues for lack of extensive overlapping in some of the traits at least. "Ability to Learn" (I) is quite distinct from "Industry" (II) for example, as shown by May's study of the relationship between Hours of Study, Intelligence, and Scholarship (3). The correlation between "Hours of Study" and "Intelligence" was —0.35. The number of hours spent in study is one measure of industriousness, at least as defined in the rating scale. The definition of "Industry" is in terms of "power of concentration and sustained attention as shown by perseverance and application to studies." Hours of study, then, should be highly correlated with "Industry," and yet the correlation obtained between "Industry" (III) and "Ability to Learn" (I) is positive. West High School teachers' ratings correlate +0.39 for boys and +0.73(!) for girls. Central High School teachers' ratings on these two traits correlate +0.64 for both boys and girls, and ratings by Humboldt High School teachers on the same traits correlate +0.55 and +0.50 for boys and girls, respectively.

Fairly high intercorrelations may be expected between some of the other traits, such, e.g., as "Ability to Learn" (I) and "Initiative" (II), "Initiative" (II) and "Dependability" (V), "Industry" (III) and "Scholastic Zeal" (IV), and "Capacity for College Work" (VIII) and all the other traits, though the precise degree of correlation which should be expected would be difficult to determine.

The ratings obtained from Central High School and Humboldt High School contained a sufficient number of ratings on Traits VI and VII ("Outside School Interests" and "Civic and Business Interests") to allow for obtaining reliable intercorrelations between these traits and the other traits.

There is no reason to expect that either of these traits should correlate with the other traits to any marked extent, since all of the other traits are aimed to differentiate good students from poor students. No one has demonstrated that a high correlation exists between extra-curricular activities or civic and business activities and scholarship. But in these ratings the average intercorrelations between Traits VI and VII are almost as high as the average intercorrelation between "Capacity for College Work" (VIII) and all of the other traits, each of which presumably is indicative of general capacity for college work.

It seems probable that, although there is some overlapping of traits in this scale, the chief factor which causes high intercorrelations is the "halo" factor. This factor affects the ratings of girls more than the ratings of boys In general, this factor has operated to reduce the reliability of these ratings as regards specific traits, through obscuring critical analysis by the presence of general attitudes In a later part of this study evidence will be presented which will enable us to define more definitely the probable basis of the general attitudes referred to.

SUMMARY

From the foregoing study of teachers' personality ratings it appears that such ratings lack high degrees of reliability, as evidenced by the following results.

1 Skewed distributions of ratings are generally obtained, showing that teachers do not differentiate equal degrees of possession of personality traits from the lowest to the highest

2 "Spread" of ratings *varies* to a marked extent from one teacher to another, though the variability of all ratings by all teachers is approximately normal

3. Although groups of teachers agree on the relative degree of possession of personality traits by their students, individual teachers vary markedly in their tendencies to rate "high," "medium," or "low," thus reducing the reliability and "comparability" of ratings by different teachers.

4 Teachers were unable to differentiate differing groups when rated one year apart In other words, they do not generally shift their rating tendencies to conform to group averages

5. Sex differences in ratings exist since both men and women teachers tend to rate girls as superior to boys, even though test records show an opposite trend This factor, however, is not very important, since differences within either sex are much more pronounced than are differences between the sexes

6. When ratings are "corrected" in order to reduce differences in *rating standards the results are not markedly different from the* original ratings.

7 The "halo" factor is present in teachers' ratings and operates to reduce their reliability This factor affects the ratings of girls to a greater extent than the ratings of boys

It appears that there are a number of factors at work to reduce the reliability of teachers' ratings of personality when those ratings are obtained under usual conditions as they exist in large high schools. In order to increase the reliability of such estimates, careful training of teachers in habits of observation and study of individual differences among their pupils seems desirable and essential if we are to obtain valid measures of those personality traits for which there are now no objective tests

REFERENCES

1 CUNNINGHAM, A B The sex element in college grades *School & Soc* 1921, **14**, 523-525

2 HART, H., & OLANDER, E Sex differences in character as indicated by teachers' ratings *School & Soc*, 1924, **20**, 381-382

3 MAY, M Predicting academic success *J Educ Psychol.*, 1923, **14**, 429-440

4 PATERSON, D G The Scott Company graphic rating scale *J Person Res*, 1922, **1**, 361-376

5 PATERSON, D G, & LANGLIE, T A The influence of sex on scholarship rating *Educ Admin & Super*, 1926, **12**, 458-469

Wesleyan University
Middletown, Connecticut

SCHOLASTIC ACHIEVEMENT OF INDIAN PUPILS*[1]

University of Oklahoma

HARVEY C. HANSEN

The American people, in assuming the overlordship of the western continents, fell heir to recurrent racial problems. The early attempts at dispossession of the native inhabitants culminated in the "Century of Dishonor." Although this period has been succeeded by several decades of altruistic endeavor on behalf of the Indians, it is generally conceded that the natives are not yet thoroughly absorbed into the American social fabric. Scientific investigators and lay friends of the Indian have advanced a number of reasons in attempting to account for this fact. Among those most frequently mentioned is the prevalence of segregated education of Indians.

A departure of the last few years that has drawn much favorable comment is the popular trend toward the inclusion of Indian children in the public school system. In Oklahoma, where approximately one-third of the nation's Indians live, 85 per cent of the Indian pupils are now enrolled in the public schools. This changed situation has given rise to questions concerning the success with which white and Indian children are being educated together in these schools.

To provide a partial answer to these questions, certain aspects of the comparative achievement of 529 Indian children enrolled in public schools, 507 of their white classmates, and 516 Indian boarding school students were investigated. Children selected from 34 public schools and seven boarding schools were subjected to achievement tests. To make the public school Indian group as representative as possible, city, village, consolidated, and one-room rural schools were included in the testing. The study covered a major portion of those sections of Oklahoma which have any considerable number of

---

*Received in the Editorial Office on December 11, 1936.

[1] A report of a field study on which was based a thesis presented for the degree of Doctor of Education at the University of Oklahoma in 1935. Acknowledgment is made of the guidance of Dr. John F. Bender, who directed the study, and of the advice of Dr. Henry D. Rinsland, who aided in the statistical treatment of the data.

Indian children in the public schools The larger public schools at which tests were given included those at Pryor, Pawnee, Hominy, Pawhuska, Stilwell, Tahlequah, Castle, and Anadarko, and the training school at Northeastern State Teachers College

At scores of schools the extent of Indian enrollment was investigated Whenever a public school with a number of Indians in attendance was visited, all of them were tested. At the same time approximately equal numbers of white pupils enrolled in the same grades were examined Those selected very adequately represented their respective racial types

The Indians were representative also from the tribal standpoint, for they came from 21 tribes

To facilitate closer analysis, the data concerning Indian pupils of half blood were separated from those pertaining to full blood Indians Table 1 gives the number and the mean age of the pupils tested in each school-racial group.

TABLE 1

NUMBER TESTED AND MEAN AGE FOR FIVE SCHOOL-RACIAL GROUPS

	White pupils	Half blood Indians Public school	Half blood Indians Boarding school	Full blood Indians Public school	Full blood Indians Boarding school
Number	507	145	93	384	423
Mean age	9.5	10 2	11 9	10 7	12 1

TESTING MATERIAL

The seventh, or April, examination of the Oklahoma edition of the *Tracy Short Answer Tests* was used (10). Each examination provides from four to seven subject tests for each grade There were 8266 separate subject tests scored

The Tracy Tests were chosen in preference to any of the published achievement tests which are nationally standardized because the former have a curricular validity in Oklahoma that probably is not approximated by any other published test that could be used for all elementary grades The nationally standardized tests base their norms on the average performance of thousands of subjects who have used a great variety of textbooks, but they are relatively invalid for testing pupils who have used textbooks adopted in a given state. In the Tracy Tests, on the other hand, the questions are based on the prescribed textbook and state course of study material covered during the preceding four weeks The teachers' compliance with state requirements is not perfect, but variations from school to school affect white and Indian pupils' test scores to approximately the same degree.

The validity of the Tracy Tests for the particular aim of this study was enhanced because their use minimized the effect of those differences in ability between Indian and white pupils which are due to general environmental factors The use of nationally standardized tests would have given the white pupil, with his wider exposure, a better opportunity to make a high grade merely because of his wider knowledge That is, his extra-curricular, and especially his extra-textbook, learning would have been capitalized to a greater extent if he had been answering the questions of a test not so closely allied with the textbooks in use in the schoolroom. The Indian child depends for a greater proportion of his exposure upon actual schoolroom situations, including the specific subject matter of the textbooks he is using Therefore, more nearly perfect curricular validity of the testing instruments operates to place the Indian child more nearly on a par with his white classmate

COMPARATIVE ACHIEVEMENT OF FIVE SCHOOL-RACIAL GROUPS

In Table 2 is presented a summary of data secured The scores for the various grades have been averaged for each subject and for each of the five school-racial groups

The white pupils, although the youngest, with an average age of 9 5 years, secured a mean score of 11 2, the public school half bloods, eight months older, a score of 10 7, the public school full bloods, fourteen months older than the white pupils, a score of 9 2.

The boarding school pupils, two and one-half years older than the white pupils and one and one-half years older than the public school Indians, secured scores of 10.0 for the half bloods and 9 3 for the full bloods.

Considered subject by subject, it is found that the general trend of the group averages of the composite mean scores is followed more or less closely by the individual subject scores Reading is a partial exception, as in this subject the public school half blood Indians scored higher than the white pupils. Writing presents an interesting varia-

TABLE 2

COMPOSITE MEAN SCORE BY SUBJECT, GROUP AVERAGE, STANDARD DEVIATION OF GROUP AVERAGE, AND COEFFICIENT OF VARIATION, BY GROUP

	Composite mean score				
	White pupils	Half blood Indians		Full blood Indians	
Subject		Public school	Boarding school	Public school	Boarding school
Spelling	14.4	14.1	11.5	11.8	10.9
Writing	10.2	11.1	11.5	10.5	10.9
Reading	8.9	9.1	7.8	6.1	6.5
Arithmetic	9.7	8.0	9.6	7.1	8.6
Language	11.5	10.2	8.6	9.0	8.1
Social science	11.3	9.7	9.3	9.5	9.0
Physiology	15.3	14.2	11.4	12.2	11.8
Group average	11.2	10.7	10.0	9.2	9.3
Standard deviation	7.0	6.7	6.3	6.9	7.0
Coefficient of variation	62.5	62.6	63.0	75.0	78.7

tion, for each one of the Indian groups tests higher than the white group, as scored on the nationally standardized Practical Handwriting Scale (11).

The coefficients of variation for the two full blood groups are considerably higher than those for the other groups. This situation is a natural concomitant of the overageness of many of the full blood Indians, as the retardation of a portion of the group causes the scores to have a wider range.

COMPARISON OF THREE PUBLIC SCHOOL RACIAL GROUPS

Because state adopted texts were not so carefully followed in the boarding schools as in the public schools, and also because the pupils in the boarding schools were taught by different teachers and under different instructional surroundings, the boarding school data do not have the validity that the public school findings possess. Therefore, to permit a more exact statistical treatment of the differences of the means, the scores of the three public school groups were compared with each other, with the boarding school data eliminated from consideration.

Table 3 gives the mean age of each of the three public school groups.

The number of tests administered in each group, enumerated subject by subject, is given in Table 4. The composite mean scores and the standard deviations—data needed in calculating the stand-

TABLE 3

MEAN AGE FOR THREE PUBLIC SCHOOL GROUPS

	White pupils	Half blood Indians	Full blood Indians
Mean age	9.5	10.2	10.7

TABLE 4

NUMBER OF TESTS FOR THREE PUBLIC SCHOOL GROUPS BY SUBJECT AND GROUP

	Number of tests		
Subject	White pupils	Half blood Indians	Full blood Indians
Spelling	505	145	382
Writing	505	145	376
Reading	507	145	378
Arithmetic	507	145	382
Language	285	98	194
Social science	218	74	118
Physiology	164	51	117
Total	2691	806	1977

TABLE 5

COMPOSITE MEAN SCORE AND STANDARD DEVIATION FOR THREE PUBLIC SCHOOL GROUPS BY SUBJECT AND GROUP

	White pupils		Half blood Indians		Full blood Indians	
Subject	Composite mean score	S D	Composite mean score	S D	Composite mean score	S D
Spelling	14.4	8.4	14.1	7.6	11.8	7.8
Writing	10.2	7.7	11.1	7.7	10.5	8.4
Reading	8.9	5.4	9.1	3.0	6.1	5.2
Arithmetic	9.7	6.4	8.0	6.3	7.1	6.2
Language	11.5	5.8	10.2	5.3	9.0	5.3
Social science	11.3	5.8	9.7	5.3	9.5	3.9
Physiology	15.3	5.1	14.2	5.0	12.2	5.4
Group average	11.2	7.0	10.7	6.7	9.2	6.9

ard errors of the means and thence the standard errors of the differences—are shown in Table 5.

Table 6 shows the algebraic differences between the respective mean scores of the white pupils and of the public school half blood Indians, together with the significance of the differences expressed as chances in 100 of true differences.

It is to be noted that when the "difference" column carries a minus quantity, in the "chances" column the corresponding number indicating chances in 100 is underlined. Such an underlined number

TABLE 6

DIFFERENCE AND SIGNIFICANCE OF DIFFERENCE BETWEEN WHITE PUPILS' COMPOSITE MEAN SUBJECT SCORE AND PUBLIC SCHOOL HALF BLOOD INDIANS' COMPOSITE MEAN SUBJECT SCORE

	Composite mean subject score			
Subject	Difference (White minus half blood)	Standard error of difference	Difference divided by S D of difference	Chances in 100 of true difference
Spelling	3	7	.4	65
Writing	— 9	7	—1 3	90*
Reading	— .2	4	— .5	69
Arithmetic	1.7	6	2 8	100
Language	1 3	.6	2 2	99
Social science	1 6	7	2 3	99
Physiology	1 1	8	1 4	92

*An underlined number is to be read as chances in 100 that the true difference will favor the second—that is, the half blood group

indicates the chances in 100 that the true difference will favor the second—that is, the Indian—group.

Accordingly, the table indicates that there are 65 chances in 100 that white children in Oklahoma schools will receive better spelling scores than will their half blood classmates, but there are 90 chances in 100 that the superiority will be reversed in the case of writing. There is a practical certainty that scores in arithmetic, language, and social science will favor white children On the other hand, there is better than an even chance that the half blood Indians, with their eight additional months of chronological age, will secure higher reading scores than the younger white children.

In Table 7 is found a comparison between public school white and public school full blood Indian scores The white pupils' scores significantly exceed the full bloods' scores in every case except in writing, where the difference is reversed

Table 8 reveals the degree of overlapping that is found when the white pupils' distribution is compared with that of each of the Indian groups The difference between the mean score of the white group and the mean score of the half blood group is 5 This is divided by the standard deviation of the scores of the half blood group, or 6 7, which gives the quotient 07 This represents the difference between the means in terms of the standard deviation. Then a standard table of areas of the normal probability curve is

TABLE 7

Difference and Significance of Difference between White Pupils' Composite Mean Subject Score and Public School Full Blood Indians' Composite Mean Subject Score

	Composite mean subject score			
Subject	Difference (white minus full blood)	Standard error of difference	Difference divided by S D of difference	Chances in 100 of true difference
Spelling	2 6	6	4.3	100
Writing	— .3	5	— 6	73*
Reading	2 8	4	7 0	100
Arithmetic	2 6	4	6 5	100
Language	2 5	5	5 0	100
Social science	1 8	5	3 6	100
Physiology	3 1	6	5 2	100

*An underlined number is to be read as chances in 100 that the true difference will favor the second—that is, the full blood group

TABLE 8

Percentage of Half Blood and of Full Blood Public School Indians Whose Individual Mean Scores Equal or Exceed Composite Mean Score of White Group

	White and half blood groups	White and full blood groups
Composite mean score	White 11 2 Half 10 7	White 11 2 Full 9 2
Difference between means	5	2 0
Standard deviation of composite mean score of Indian group	6.7	6 9
Difference between two means in terms of standard deviation	07	29
Difference between two means in area of normal distribution	2 8%	11 4%
Percentage of Indian group equaling or exceeding mean of white group	47 2%	38 6%

entered in terms of the standard deviation, this procedure yielding the difference between the means as a percentage of the half blood normal distribution If this 2 8 per cent of the Indian group be subtracted from the 50 per cent that lies above its own mean, the resulting remainder indicates that 47 2 per cent of the Indian group lies above the mean of the white group

By the same method of computation it is found that 38 6 per cent of the individual full blood Indians reach or exceed the average performance of the white pupils

APPRAISAL AND CONCLUSIONS

There are certain difficulties which are frequently encountered in the carrying out and interpretation of comparative racial studies of school achievement Races are in a state of flux, there is racial feeling, either of antagonism or of inferiority, language handicaps, especially in the case of bi-lingual subjects, affect the results, irregular attendance conceals potential achievement; the incentive to learn may be lacking In short, it is frequently declared that the entire pattern of home and community environment is such that no attempt can be made to equate racial groups (1, 2, 5, 9, 12)

While acknowledging the truth of the principles stated above, it should be noted, in appraising the results of this study in so far as they relate to Oklahoma public school pupils, that the Indian child is taken as he is, as a matter of choice He is tested in the setting in which he is found, both because that is where he can the most easily be tested and also because the purpose of the investigation is to ascertain how successfully he and his white classmates are being educated together Whatever the cause may be, if the Indian's test score is low, so also is his grade of achievement in his ordinary schoolroom work The examination is a real measure of his usual scholastic success

The data presented appear to justify the following conclusions

1 No clear-cut difference in achievement between public and boarding school pupils of Indian blood is revealed (8) The public schools seem to serve the academic needs of the younger, more able Indian pupils, the boarding schools, the requirements of the older, more retarded pupils

2 In the public schools there is a very definite superiority in the general achievement of the white child as compared to that of the half blood Indian A greater difference is found when the white child's scholastic success is compared with the full blood's In every subject except writing the white child excels, with the differences statistically reliable This situation, combined with the disparity in ages, introduces an element of heterogeneity into the public schools that may have a detrimental effect upon instruction

3 The data secured in this investigation are entirely consonant

with the results of many earlier studies which have been reviewed, for the secondary data show, almost invariably, a high percentage of retardation (8), and a low level of measured intelligence of Indian pupils (3, 4, 6, 7). This intellectual and school achievement status is due to the entire socio-racial pattern of influences which bear upon Indian pupils. The reality and relative weight of those factors which may be biologically innate remain undetermined.

4. The marked variability of scores within the full blood groups and the high percentage of overlapping of white scores by Indian scores is suggestive. While the majority of the Indian pupils tend to lower the average performance of the grades and schools of which they are members, a few individuals among them do excellent work and present no administrative problem. They have possibilities as future leaders and citizens.

REFERENCES

1. BLACKMAR, F. W. The socialization of the American Indian. *Amer. J. Sociol.*, 1929, **34**, 653-669.
2. DANIEL, R. P. Basic considerations for valid interpretation of experimental studies pertaining to racial differences. *J. Educ. Psychol.*, 1932, **23**, 15-27.
3. FITZGERALD, J. A., & LUDEMAN, W. W. The intelligence of Indian children. *J. Comp. Psychol.*, 1926, **6**, 319-328.
4. GARTH, T. R. A comparison of mental abilities of mixed and full-blood Indians on a basis of education. *Psychol. Rev.*, 1922, **29**, 221-236.
5. GARTH, T. R., SERAFINI, T. J., & DUTTON, D. The intelligence of full-blood Indians. *J. Appl. Psychol.*, 1925, **9**, 382-389.
6. GARTH, T. R. The intelligence of mixed-blood Indians. *J. Appl. Psychol.*, 1927, **11**, 268-275.
7. GARTH, T. R., SMITH, H. W., & ABELL, W. A study of the intelligence and achievement of full-blood Indians. *J. Appl. Psychol.*, 1928, **12**, 511-516.
8. GARTH, T. R., & GARRETT, J. E. A comparative study of the intelligence of Indians in United States Indian schools and in the public or common schools. *School & Soc.*, 1928, **27**, 178-184.
9. JAMIESON, E., & SANDIFORD, P. Mental capacity of South Ontario Indians. *J. Educ. Psychol.*, 1928, **19**, 313-328.
10. RINSLAND, H. D. (ED.) Tracy short answer tests. Oklahoma edition. Kansas City: Dodsworth.
11. ——— Practical handwriting scale. Dallas: Practical Drawing Co., 1930.
12. YODER, D. Present status of the question of racial differences. *J. Educ. Psychol.*, 1928, **19**, 463-470.

College of the Ozarks
Clarksville, Arkansas

THE DEVELOPMENT OF VISUAL ACUITY IN THE CAT*

University of Rochester

J WARKENTIN AND K U. SMITH

I INTRODUCTION

In recent years the fact has become increasingly evident that the complex nature of vision in the adult mammal must be understood partly in terms of the growth and development of visual capacity in ontogeny. Animals of the same species vary greatly at different ages in their capacity to respond to visual and other types of external stimulation, and these differences may be shown to be in part a function of the past history and development of the individual. The character of vision in the adult mammal thus represents an integration of physiological processes which have passed through various stages of development in the life history of the organism The description of these stages of development of visual capacity and associated physiological changes constitutes one main aspect of the general problem of vision

Although numerous studies have been made of the ontogenetic development of visually determined responses in animals, the relation between the appearance of such reactions and the sensory capacity of the organism is not yet clear. This lack of knowledge may be attributed largely to the absence of methods which will permit quantitative measurements of visual capacity during the early period of growth In the present study a method is described which makes possible the quantitative determination of visual acuity in young mammals and results are presented concerning the relation between visual acuity and the development of certain visually determined reactions in the cat

II PRIOR STUDIES OF THE ONTOGENY OF THE VISUAL MECHANISM

Several approaches have been utilized in the attempt to ascertain the character of visual development in different animals. Usually

---

*Accepted for publication by Leonard Carmichael of the Editorial Board, and received in the Editorial Office on December 15, 1936

these techniques were not originated specially for use in a study of ontogenetic development, but rather in phylogenetic studies. The techniques used up to the present time in such studies may be grouped under three headings. (*a*) Experiments centering about the observation of the development of various responses during the process of growth, (*b*) studies of the electrical excitability of the cortex and other nervous structures at different ages, (*c*) investigations of the degree and sequence of myelination of tracts in the nervous system, especially that part of the nervous system related to the optic connections.

Table 1 summarizes some of the experimental results bearing upon the development of visually controlled behavior and other growth changes associated with vision in normal animals Although incontestable conclusions concerning the development of vision in ontogeny cannot be derived from these data, they emphasize further the extensive variations occurring in animals at different stages of growth A more detailed account of the literature related to the present problem is presented in the following paragraphs

1. *Behavioral studies.* The investigation of the ontogeny of vision consists in the determination of the sequences of growth which lead to the appearance of visually controlled behavior. It is necessary to realize that the word "vision" is used to characterize a number of different types of behavioral adjustments, ranging from closure of the pupils to the highly verbalized perceptual reactions of the adult human individual. The study of the ontogeny of vision therefore entails a consideration of the development of the many different aspects of behavior which are subject to visual control

Beginning with reports regarding the ontogeny of vision in the lower mammals, it is found that Small (37) observed that young albino rats open their eyes at the ages of 15 to 17 days, give lid reflexes to light at the age of 17 days, and only at a later time respond to moving objects by means of visual pursuit reactions Tilney (42) observed that the eyelids open at the age of 14 to 17 days in the rat Visual pursuit reactions of the head and eyes occurred as soon as the eyes were fully open, but at this age vision did not seem to be involved in the general behavioral adjustments of the animals Turner (45) studied the development of vision in rats by means of a technique in which the animals were required to discriminate the position of a black exit door in an otherwise

TABLE 1

THE AGE IN DAYS WHEN EYE OPENING AND CERTAIN VISUALLY DETERMINED REACTIONS WERE FIRST OBSERVED BY DIFFERENT EXPERIMENTERS ON VARIOUS ANIMALS

Species	No of animals	Eyes open	Pupil resp	Lid resp	Visual pursuit	Air right	Experimenter & date	
White rat	3-10	15-17		17			Small	1899
Guinea pig	3	0*			2		Mills	1898
Guinea pig	ca 32	0	0	0			Avery	1928
G-pig fetus	2-8	55-59	59				Avery	1928
G-pig fetus	2-4	56	41-45	61			Carmichael	1934
Rabbit	2	10-12			22		Mills	1898
Rabbit	4	8				18	Magnus	1924
Rabbit	71	12	12				Kao (Shirley)	1927
Dog, St Bern	7	13		15	18-23		Mills	1898
Dog, mongrel	7	10		20	20		Mills	1898
Dog	1-18	16	5-7**				Michailow	1910
Cat	1	10	12	11	14		Mills	1898
Cat	?	7	7		14		Tilney & Casamajor	1924
Cat	2-9?	6-10			14		Langworthy	1929
Cat	2-59	7-10	7-10	0	15-21	22-28	Windle	1930
Monkey	3	0	2-3	8	5-11		Tinklepaugh & Hartman	1932
Chimpanzee	1	0	1-7	1	7?		Jacobsen, Jacobsen & Yoshioka	1932
Human infant	6	0			1-14		McGinnis	1930

*In this table, a "0" indicates that the response was present before the end of the first day

**The eyelids were removed surgically at early ages brightly lighted field. The results of this experiment indicate a *definite difference* between the ability of young rats to find the exit door before and after the eyes open Even if the eyes opened early, there was evidence of inferior distance vision up to the age of 16 or 17 days

Wesley Mills (28) carried out some of the first observations on the ontogeny of vision in neonatal animals In observations of 3 guinea pigs he found that the eyes were open at birth, and that the lid reflex to light was present also at that time Visual pursuit reactions were observed in these animals when they were 2 days old Some years later, Avery (3), in a study of several late guinea pig fetuses, found that the eyes were opened at gestation ages between 55 and 59 days, that the pupil response to light was present at 59 days, and that visually determined avoidance of obstacles took place at 63 days of age Newborn guinea pigs had their eyes open and gave pupil and lid response to light. In a still more recent study of fetal guinea pigs, Carmichael (7) found that, if the eyelids were surgically opened, eyeball movements correlated with light stimulation were observed after the gestation age of 45 days Responses of the iris to light were observed at 41 to 45 days gestation age

In a study of two young rabbits, Mills (28) found that the animals opened their eyes between the ages of 10 and 12 days, and gave visual pursuit reactions at the age of 22 days. Magnus (24) found the eyes open at the age of 8 days in 4 rabbits. He also observed that air righting reactions, which are in part visually determined (Muller and Weed, 30), were present at the age of 18 days. Kao (36) observed the pupil response to light in rabbits on the same day that the eyes opened, i e., at the age of 12 days.

Mills (28) also investigated the development of 7 St Bernard puppies and of 7 mongrel puppies The former opened their eyes at the age of 13 days, and visual pursuit movements appeared at the ages of 18 to 23 days The mongrel puppies opened their eyes at the age of 10 days, gave the lid reflex to light at 20 days, and visual pursuit movements also at the age of 20 days Michailow (27) found that pupil contraction to light in dogs with operatively removed eyelids was present at ages between 5 and 7 days A kitten studied by Mills (28) opened its eyes at the age of 10 days, gave lid reflexes to light at the age of 11 days, the pupil reflex at the age of 12 days, and visual pursuit movements at the age of 14 days

Tilney and Casamajor (43) state that the eyes open on the seventh day in the kitten, that the pupil response is present at that same age, but that visual pursuit reactions could not be elicited until the age of 14 days Langworthy (20), observing several kittens, found that the eyes opened at ages between 6 and 10 days. At the age of 14 days the kittens gave visual pursuit reactions, while the air righting reflex was developed during the fourth week of life In a large series of kittens, Windle (49) observed that the eyes opened at ages between 7 and 10 days, at which time the pupil response was also present Contrary to the observation of Mills, he found that the lid reflex to a bright light was active at birth At 14 days of age, Windle's kittens avoided objects visually, at 15 to 21 days they gave visual pursuit reactions, and during the fourth week they developed the falling reflex

Two studies are available on the development of visually determined responses of infra-human primates. Tinklepaugh and Hartman (44), in observations on 3 monkeys, determined that the pupil reflex to light was present at the ages of 2 to 3 days The lid response was observed only on one monkey, and then at the age of 8 days. Visual pursuit movements could be elicited at ages of 5 to 11 days, but such responses could never be observed if a striated, rotating visual field was presented to the animals. This latter aspect of the experiment of Tinklepaugh and Hartman seems difficult to explain in the light of the results of a similar experiment by McGinnis (25) on the human infant, in which it was found that oculocephalogyric responses were induced by rotating the visual field Another experiment is that of Jacobsen, Jacobsen, and Yoshioka (15), who observed an infant chimpanzee from the time of birth It was found that the lid reflex could be elicited during the first day Pupillary constriction to bright light was also seen on the first day. Although this reaction was quite weak at such an early age, it became more and more definite during the first week of life Visual pursuit movements were seen at the age of about one week

A large number of observations of the development of visually determined responses have been reported in studies of the human infant, but the majority of these studies were not carried out under controlled experimental conditions and the results are consequently of qualitative nature The literature pertaining to studies of the appearance of the pupillary and lid reflexes, fixation, conjugate eye movements, visually determined grasping reactions, and other types of visually determined behavior have been summarized by Jones (16), Peiper (31), McGinnis (25), Shirley (36), and Pratt (32)

The experiment of McGinnis (25) has already been mentioned It is of special interest because this author was able to elicit oculocephalogyric responses in human infants during the first 12 hours after birth Although these responses seemed to be weak and infrequent at this time, they became more definite as well as more frequent with increasing age Successful ocular pursuit movements were observed by the age of 2 weeks, and by the age of 6 weeks these reactions were quite regularly elicited

2. *Electrical Stimulation Experiments* Although studies on the electrical excitability of the cortex and of other neural structures constitute what may be considered as an indirect approach to the problem of the appearance of visual function, they shed valuable light on the neural mechanism involved in the development of visually determined reactions.

One of the earliest studies of the electrical excitability of the cortex of young mammals is that of Soltmann in 1876 (41) He seemed to find a correlation in time between the opening of the eyes in young dogs at the age of 8 days and the first responses of the forelimbs which could be elicited by cortical stimulation at the age of 10 days. More recently, Michailow (26) performed a number of careful observations on the electrical excitability of the cortex in young dogs. In puppies less than 24 hours old this investigator was able to elicit turning of the head in a horizontal plane, contractions of the neck muscles, and raising of the head At the age of 3 days, electrical stimulation of the cortex elicited lateral movements of the eyes, which were often combined with movements in a vertical plane Two days later, Michailow saw turning of the head about the body axis and closing of the eyelids. At the age of 9 days stimulation elicited enlargement of the pupils, while conjugate movements of the eyes appeared at the age of 10 days This study appears to be one of the first reported in the literature which presents an analysis of the cortical centers from which eye reflexes and head movements could be elicited in young mammals Michailow (27) also carried out a study of the pupillary reflex in 18 puppies After removing both upper and lower eyelids in the young dogs, Michailow exposed the vagosympathetic nerve on one side of the neck, and stimulated the nerve electrically He found no evidence of normal eye reflexes at birth, but observed that these developed at various ages after birth The nictitating membrane reflex was the first eye reflex to be elicited by stimulation of the nerve, and was present at the age of 3 days After the age of 5 days the pupils began to respond to light, but stimulation of the nerve did not elicit pupil responses till the age of 7 days At the age of 7 to 8 days, forward movements of the eyeball could be elicited by stimulation of the vagus nerve Weed and Langworthy (47, 48) studied the electrical excitability of the cortex in the opossum and in the cat In the opossum they found responses of the forelegs at the age of 23 days, long before the eyes opened Eye movements could not be cortically elicited before the age of 76 days, after the eyes had been open for some time. In the study of the cat a similar result was secured, in that electrical excitation of the cortex elicited foreleg responses at birth, but the eyes did not open until the animals had reached an age of from 7 to 9 days

3 *Myelinization Studies* Although the myelinization of nervous tracts has never been fully accepted as a criterion of nervous function, it has nevertheless been correlated in several experiments with stages of visual development in young animals. In 1896 Held (14) attempted to prove his theory that the degree of myelinization is an index of nervous function and that conversely, impulses passing over nerve tracts exert a myelin-forming influence on such tracts. Held opened the right eyes of rabbits, cats, and dogs soon after birth, exposed these animals with one eye open to light for periods of time varying from 5 to 60 hours, and then examined the degree of myelinization of the optic nerves In each animal examined he found that the right optic nerves showed a more advanced degree of myelinization in comparison with the left optic nerves. Results of somewhat similar significance were obtained by Berger (6) He sutured the eyelids of 3 dogs and 4 cats, and found that, when these animals were killed and their central nervous systems examined at ages varying from one to 10 months, the gyri and cells of the cortex which are involved in visual responses were definitely underdeveloped

Although such investigators as Allen (1), in her work with guinea pigs, and Angulo y Gonzales (2), in his work with rats, have found functional activity in nerve tracts before they were myelinated, thus proving that myelinogeny is not an absolute index of behavioral capacity, findings to the contrary must also be taken into consideration. A much quoted experiment directly related to the present study is that of Tilney and Casamajor (43) on the development of behavior in kittens These authors observed the development of kittens from birth till the age of about one month. After observation of some new increment of behavioral development the central nervous system was examined for the relative degree of myelinization It has been noted above that they found that the eyes opened during the seventh day in the kitten, an observation which fails to agree with the ages of opening of the eyes noted by other investigators The number of animals used is not stated Tilney and Casamajor conclude that the explanation of the eye opening reaction is to be seen in the myelinization of fibers in the optic tracts. On or about the sixth day they found the fibers constituting the optic chiasm and tract fairly well myelinated. They found further that only pupillary fibers are myelinated all the way from the retina to the oculomotor nucleus at this age, a fact which is interpreted by Tilney and Casamajor to be the explanation of the appearance of the pupillary and lid reflexes at the time when the eyes are opened No other visually determined behavior was seen until the myelinization of the optic fibers up to the superior colliculus, lateral geniculate body and the pulvinar at the age of 14 days. At this time the first oculocephalogyric reactions were elicited, but only in the horizontal direction Vertical pursuit movements were seen on the following day

Langworthy has carried out several experiments relating to the correlation between myelinization and the appearance of neural function. This author (18) found all cranial nerves of 3 kittens myelinated during the first day after birth In a study of the young opossum, Langworthy (19) found complex activity in this animal at the time of birth, a time before there was any evidence of myelin formation on nerve tracts In the opossum, fibers connected with the oculomotor nerve first showed traces of myelin at the age of 41 days By the age of 46 days all cranial nerves showed evidence of some myelin, but at 56 days there was still no myelin present on the optic tracts. At the age of 62 days one animal had its eyes open, and the optic tracts were beginning to myelinate up to the superior colliculus At the age of 64 days optic fibers were definitely myelinated to the lateral geniculate body At 76 days it was observed that the nerves of the extra-ocular eye muscles had a few myelinated fibers Langworthy concluded that the opossum is capable of complicated reflex activity before any myelinated fibers are present, and that the correlation between the development of function and myelinization is certainly suggestive, but by no means positive In a further study of kittens, Langworthy (20) corroborated the findings of Tilney and Casamajor (43), that the establishment of reflex connections of myelinated optic fibers with the oculomotor nucleus brings about the opening of the eyes. Langworthy observed that the eyes opened at the ages of 6 to 10 days, that visual pursuit responses occur first at the age of 14 days, and that air righting develops between the ages of 21 and 27 days The number of kittens used in this study is not given, nor is Langworthy's evidence on the physiological basis of the opening of the eyes entirely clear

The experiment of Fish and Windle (8) on the effect of rotatory stimulation of the movements of the head and eyes in young kittens does not relate directly to the present review, but the conclusion of these authors regarding the effect of surgical opening of the eyelids is of interest In 31 kittens they opened one or both eyes before the age of 3 days and found that the appearance of ocular movements seemed to be precocious as a result of a very early opening of the lids

Two additional recent studies will end this review Goodman (10) kept 9 rabbits of one litter in a perfectly dark room from birth until the age of 5 to 6 months When the animals were first exposed to light stimulation, they gave evidence of good lid and pupil reflexes Goodman concluded that complete lack of optic stimulation in rabbits from birth to the age of 5 or 6 months did not significantly retard either the functional capacity or the degree of myelinization of the optic system of these animals However, one of the animals used by Goodman required approximately 2 weeks in order to develop normal visually determined reactions In the same connection the study of Mowrer (29) on pigeons is significant Young pigeons were raised with the lids sutured till the age of 6 weeks When exposed to light at this age, the animals gave good pupillary reactions to light, but showed only very feeble optic nystagmus and no visually determined avoidance reactions. Normal avoidance and nystagmus responses appeared within 3 days after the eyes were opened

From the foregoing review of these studies it is seen that the fact is not yet proved that myelinization is an infallible index of the beginning of nervous function, and that the use of myelinization as a criterion for judging the beginning of function may be open to some question.

III. The Present Investigation of the Development of Visual Acuity in the Cat

1 *Problems.* The best conclusion which can be reached from a consideration of studies reviewed in the last section is that, in different species of animals, various visually determined reactions appear at different ages, and that the appearance of these reactions may be correlated to some extent with such physiological factors as myelinization and cortical excitability However, none of the investigators cited and none to the writer's knowledge have made an attempt to investigate the relation between the development of visually controlled behavior and the measured sensory capacity of the animal at the time of observation Consequently, it is not known whether the development of vision in ontogeny consists only in the addition and integration of newly acquired patterns of response, as many investigators seem to believe, or whether such growth involves the gradual increase in acuity of vision in responding to differences in light intensity, visual patterns, and wave length. Furthermore, the question has also remained unanswered as to the degree of acuity of vision associated with the development of different visually determined reactions

In the present study experimental results are described which bear directly upon a number of such unsolved problems. Specifically, the following questions have been considered: (1) At what age in the cat can visually controlled oculocephalogyric responses be elicited? (2) What is the visual acuity of the cat, measured by means of the oculocephalogyric response, at the time of the onset of these reactions? (3) Does the visual acuity of the animal improve gradually in time from the date of the onset of visually controlled behavior? (4) What relation exists between opening of the eyes and the development of visual acuity in the cat? (5) What relation exists between visual acuity and the date of appearance of visually controlled placing responses of the forelimbs in the cat? (6) How extended is the period of the development of these phases of visually determined behavior in the young kitten?

2. *Apparatus and Technique* In the present investigation a method was employed which involved the observation of compensatory eye and head movements or oculocephalogyric reactions in response to a rotating visual field Although various modifications of this method have been employed previously in the study of visual capacity in invertebrates (Radl, 33, Schlieper, 35, Hecht and Wolf, 13, Saelzle, 34, Hecht and Wald, 12, Wolf, 50, Wolf and Zerrahn-Wolf, 51), and vertebrates (Lyon, 23, Garrey, 9; Loeb, 22, McGinnis, 25; Grundfest, 11, Tinklepaugh and Hartmann, 44, Wolf and Zerrahn-Wolf, 52), the present experiment constitutes the first attempt to make use of the principle of this technique for obtaining quantitative measures of visual capacity in young animals

A diagram and a detailed description of the apparatus used in this study has been previously presented by Smith (39) A cylinder, 105 cm. high and 120 cm in diameter, was constructed so that it could be rotated on its base independently of the animal, when the latter was placed on a small stationary platform on the inside of the cylinder The cylinder consisted of a wooden framework that was lined on its side and bottom with white oil-cloth, on the inner side of which black strips of cardboard could be inserted, varying in width from 5 cm to 0 6 cm Smaller stripes were obtained by means of cloth printed with alternate black and white lines, 0 16 cm in width The base of the cylinder was of rigid construction and contained a vertical shaft at its center This shaft was inserted in a bearing which was mounted upon a heavy iron base below the cylinder The cylinder could be easily rotated in this bearing by hand or by means of a pulley system led from the shaft of the cylinder to a motor

The animal platform, 32 cm in diameter, was suspended from the ceiling by means of an iron support rod The position of the platform could be adjusted so that the animal's head was located at a point at the center of rotation of the cylinder. The young animals were prevented from falling off of the platform by means of a fence, about 10 cm. high, that was made of very thin white silk thread wound on 8 small metal posts around the edge of the platform As far as could be observed, the thread and posts did not in any way interfere with the responses of the animal's to the rotating pattern

The opening of the cylinder, through which the animals were placed on the platform, was closed during observations by means of an adjustable white cover that was suspended from the ceiling Observations were made through a small aperture in this top.

The inside of the apparatus was illuminated by two 200-watt bulbs, one fastened to the cover of the drum above the animal's head, and the other attached below the animal platform The level of brightness used in these experiments was about 11 candles per square foot, and this effective brightness did not vary over ½ log unit with the different series of striped patterns.

In this apparatus, visual acuity can be expressed in terms of the visual angle subtended by the smallest stripes to which the animal responded from its position at the center of the cylinder. The distance intervening between the sides of the cylinder and the animal's head was approximately 50 cm throughout the experiments Hence, the lines of different widths employed represent visual angles of approximately 360, 180, 43, and 11 minutes of arc respectively.

It might be objected that such thresholds secured in relation to the oculocephalogyric reaction do not represent true measures of visual acuity In this study, the fact is assumed that the term "visual acuity" represents the threshold capacity of the organism in adjusting or orienting to changes in brightness contours in the visual field Since the observations made here involved correlating a definite orientation of the head and eyes to measurable changes in brightness contours or patterns in the field of vision, it is concluded that the term "visual acuity" is properly employed in the description of the visual function under investigation

3 *Procedure.* Observations of compensatory responses of the eyes and head to the rotating striated patterns were begun at the time when an animal opened its eyes Tests were first made with the lines 5.0 cm in width, and the lines were thereafter reduced in width when definite eye or head nystagmus was observed In a few cases it was impossible to make observations with the complete series of stripes. In examining the oculocephalogyric reactions, the cylinder was rotated through a complete revolution in one direction, followed by a complete rotation in the opposite direction This procedure was carried out 5 times daily, each rotation of the cylinder constituting one of the 10 trials. The rate of rotation was approximately two or three revolutions per minute The criterion of a positive response was that the animal make a definite pursuit movement of the eyes or head in the direction of the movement of the stripes, followed

by another pursuit movement If such a response was seen only once, or was not clearly visible to the observer, it was not accepted as positive

In addition to the observation and recording of the visual acuity of the animals, a record was kept of the date of the opening of the eyes in every case Visual placing reactions of the forelimbs were tested daily in 14 of the animals, which were selected from all litters of animals used in this experiment with the exception of the first litter In testing an animal for visual placing reactions, it was grasped by the lower trunk and pelvis, leaving the head and forelimbs free, and was moved gently toward the surface of an illuminated table Positive responses were recorded if the front legs were definitely extended before the vibrissae or other parts of the body touched the surface of the table In 13 of the animals similar observations were made on placing of the forelimbs induced by contact with the vibrissae and lower jaw on the surface of the table

4 *Animals* In all, 24 animals from 9 litters were employed in the experiment These animals were given constant care, and were weighed daily to provide an index of their general condition of health In this experiment, a daily increase in weight was taken as a criterion of selection of animals for the different observations made Only healthy animals, i e , those showing a marked daily increase in weight, were used for the visual acuity observations. Conjunctivitis was avoided by daily washing of the eyes with a 20 per cent solution of Argyrol from the time that the eyes opened

5. *Results of the present experiment* Table 2 presents a summary of the genetic development of certain aspects of vision studied in the present investigation This table indicates the age in days of opening of the eyes as well as the age of appearance of different visually determined responses in the different animals For purposes of comparison with the age of development of the visual placing responses of the forelimbs, the age of appearance of similar tactually controlled placing responses is also indicated The first column to the left identifies the litter of each animal by means of capital letters Some of the animals employed belonged to different litters taken from the same female These cases are indicated by small letters following the capital letters

a *The age of eye opening* The 24 kittens used in this investigation had their eyes completely open at ages distributed at random

TABLE 2

THE AGE IN DAYS WHEN THE EYES OPENED AND WHEN THE APPEARANCE OF DIFFERENT VISUALLY AND TACTUALLY DETERMINED RESPONSES WAS OBSERVED IN KITTENS IN THIS STUDY

Litters	Date of birth	No of kitten	Age when eyes open	Tactual placing	Head nystagmus	Visual placing
Aa*	10/29/35	1	7		17	
		2	8		16	
		3	8		17	
B	1/5/36	1	4		14	27
		2	6		14	
Ca	2/28/36	1	12		12	28
		2	12		13	
		3	12		13	28
D	3/18/36	1	7		14	26**
		2	8		14	26
		3	3		14	26
Ab	4/11/36	1	4	11	14	23
		2	4	12	12	23
E	4/27/36	1	5	11	15	22
F	5/18/36	1	11	11	14	25
		2	11	13	14	
G	5/19/36	1	12	14	15	25
		2	12	15		
		3	11	13		
Cb	6/15/36	1	15	15	14***	25
		2	15	13		
		3	10	15	13	25
		4	12	16	13	
		5	14	13	14	25

*In the litter designations, the small letters "a" and "b" stand for the first and second litters of the same mother animal.

**Kittens D-1, D-2, and D-3 were not tested at ages of 23 to 25 days.

***This kitten was tested with only one eye open at this age

from 3 days to 15 days, with no apparent relation between health or weight of the animals. Upon examination of Table 2 it will be readily observed, however, that the different members of the same litters tended to open their eyes at about the same time Litters *Ca, Ab,* and *F* show this most clearly Evidently, then, the parentage of the kitten determined in part the age at which its eyes opened This is further demonstrated by the fact that the members of litters *Aa* and *Ab* opened their eyes comparatively early, while the members of litters *Ca* and *Cb* all opened their eyes relatively late No sex differences were noted as to the age when the eyes opened in the different kittens *b The first occurrence of visually determined nystagmus.* Under the present experimental conditions, the initial appearance of the cephalogyric reactions determined by light stimulation took place in the kittens at ages between 12 and 17 days Head nystagmus was used as the major criterion of response during this period, since it is very difficult to observe movements of the almost uniformly colored eyes of the kittens during the first weeks of life The comparatively late appearance of cephalogyric reactions in the case of Litter *Aa* is to be attributed partially to lack of skill in observation on the part of the observer at the beginning of the experiment In respect to these responses, as was noted in the case of the eye opening data, there is considerable homogeneity within litters as to the age when the response first appears. Litters *B*, *D*, and *F* clearly show this tendency, while the greatest variability was found in Litter *Ab*.

c The appearance of visual placing reactions Visually determined placing reactions of the forelegs occur almost invariably when the adult cat is held by the lower trunk and is moved toward some clearly visible surface Rademaker first described these reactions, and they have been more recently studied by Bard and Brooks (4), and Bard and Orias (5), and Smith (39) In the course of the present experiment data were secured which demonstrated that visual placing reactions appear at a rather constant level of growth in the developing kitten The animals used in this study first gave visual placing reactions between the ages of 22 and 28 days Reference to Table 2 will show that wherever two or more kittens of one litter were tested, visual placing reactions were observed to appear at the same age An interesting observation in connection with the appearance of these visually determined placing reactions is that, at the time when they are first elicited, the illumination of the surface and of the room in general are important factors It was found, for example, that a kitten might place visually in a dimly lighted room, but not under conditions of bright illumination. This factor was controlled in the present experiment by testing all kittens under a variety of different conditions

d The appearance of tactual placing reactions Although tactual placing responses have no direct bearing upon the development of visually determined placing reactions, they may be considered as significant precursors of visually determined placing reactions The difference between the two types of placing reactions seems to be mainly attributable to a difference in afferent control. Therefore, the difference in age at which the two types of responses first occur may be accounted for largely in terms of the delay in the growth of the central visual pathways mediating the visually determined placing responses. The animals tested in the present study gave the first tactual placing responses at ages varying from 11 to 16 days. The period of time intervening between the appearance of tactual placing responses and visual placing responses varied from 10 to 14 days. Some relation between the age of appearance of the two types of placing responses will be seen upon reference to the table. In some cases where tactual placing was observed at a relatively early age (kittens *Ab*-1, *Ab*-2, *E*-1), visual placing responses also occurred relatively early. The influence of parentage upon the age at which placing responses appeared is again suggested.

Table 3 summarizes averages and deviations of the ages at which

TABLE 3

SUMMARY AND ANALYSIS OF DATA PRESENTED IN TABLE 2
(All ages are given in terms of days)

	Age when eyes open	Tactual placing	Head nystagmus	Visual placing
Number of kittens tested	24	13	22	14
Mean ages calculated	9.1	13.4	14.0	25.5
Mean deviations of ages	3.0	1.5	0.9	1.3
Range of ages	3-15	11-16	12-17	22-28

the eyes opened and at which the different reactions described above first occurred. It will be seen from this table that the mean age at which the kittens opened their eyes is 9 days, the mean deviation 3 days, and the range 12 days. This range is significantly larger than the range of appearance noted in any of the responses observed. In the case of the initial appearance of visually determined head nystagmus, the mean age is 14 days. It is significant to note that this mean deviation of less than a day is the smallest observed in the case of any response studied in the present experiment. The appearance of the oculocephalogyric reaction seems to be entirely independent of the age at which the eyes open, as is demonstrated by the fact that the rank difference correlation for the ages at which the eyes open and the time when head nystagmus is first observed is —.06, a value which is not significantly different from zero for the number of cases used. The average age of the appearance of visual placing reactions in the kittens observed is 25 days, with a mean deviation of about one day.

e. The development of visual acuity. The technique of the rotating visual field is admirably adapted to the determination of the acuity of pattern vision in young mammals. Obviously no training technique could be used, because considerable development in vision would occur before the kittens could be trained to make any sort of visual discrimination. The present technique permits the ready examination of visually determined responses in the kitten to various visual patterns on the same day.

Table 4 presents data secured in the present investigation regarding the development of various degrees of visual acuity in the kitten at various ages. The data regarding the appearance of responses to the 5.0 cm. stripes was discussed in connection with Table 2. It will

TABLE 4

AGE IN DAYS WHEN NYSTAGMUS RESPONSES WERE FIRST OBSERVED TO STRIPES OF DIFFERENT WIDTHS

Animals		5.0 cm stripes	2.5 cm stripes	0.6 cm stripes	0.16 cm stripes
Aa	1	17			
	2	16			
	3	17			
B	1	14	17	24	
	2	14	17	24	
Ca	1	12		23	30
	2	13		23	23
	3	13		23	30
D	1	14	15	21	26*
	2	14	15	21	27
	3	14	15	22	27
Ab	1	14	19	20	23
	2	12	19	20	23
E	1	15	18	19	24
F	1	14			25
	2	14			
G	1	15			24
Cb	1	14	15	19	24
	2	13	15	20	
	3	13	15	19	24
	4	13	15		24
	5	14	15	20	24

*Kittens *D*-1, *D*-2, and *D*-3 were not tested at ages of 23 to 25 days.

be seen from Table 4 that responses to the 2.5 cm. stripes were observed in various kittens at ages between 15 and 19 days, responses to the 0.6 cm. stripes were seen at ages between 19 and 24 days, and responses to the 0.16 cm. stripes could be elicited at ages between 23 and 30 days. As in the case of the other responses discussed above, there was a marked homogeneity within litters as to the ages at which the kittens responded to the stripes of various widths. The period of time during which responses were observed to develop in the different animals to all four widths of stripes varied from 9 to 18 days, with an average of 11 days.

The development of visual acuity in the animals used in this study can be expressed in terms of the visual angles subtended by the narrowest lines responded to at different ages. The line widths employed correspond to visual angles of 360, 180, 43, and 11 minutes of arc respectively under the conditions of the experiment. Using this method of description, it may be said that, at the time of the first onset of the oculocephalogyric reactions at 14 days, the visual acuity of the kitten is very poor. From then on until the age of 4 weeks, there is a gradual increase in acuity of vision, which reaches a value of at least 11 minutes of arc between 23 and 30 days of age.

The data regarding the development of visual acuity in kittens is more briefly summarized in Table 5 in terms of averages, mean deviations, and ranges of ages at which the responses were observed. It is seen from this table that responses to the 5.0 cm. stripes were observed at an average of 14 days, that responses to the 2.5 cm. stripes usually appeared two days later, that responses to the 0.6 cm. stripes were observed at an average age of 21 days, and that 4 days after this, at an average age of 25 days, the kittens responded to the 0.16 cm. stripes. The mean deviations for the first occurrence of these

TABLE 5
SUMMARY AND ANALYSIS OF DATA PRESENTED IN TABLE 4
(All ages are given in terms of days)

	5.0 cm. stripes	2.5 cm. stripes	0.6 cm. stripes	0.16 cm. stripes
Number of kittens tested	22	13	15	15
Mean ages calculated	14.0	16.2	21.2	25.2
Mean deviations of ages	0.9	1.4	1.7	1.9
Range of ages	12-17	15-19	19-24	23-30

responses indicating different degrees of visual acuity varied from less than a day in the case of responses to the 5.0 cm. stripes to about two days in the case of the 0.16 cm. stripes. This comparatively high mean deviation of ages when responses first occurred to the 0.16 cm. stripes is to be attributed largely to the fact that two kittens, *Ca*-1 and *Ca*-3, which were difficult to observe, did not show clear-cut responses to the small lines until the age of 30 days.

It is a significant fact that the age of appearance of visual placing reactions was 25 days, and that the average age at which responses to the 0.16 cm. stripes were observed was also 25 days. That a high degree of visual acuity is correlated in time with the development of

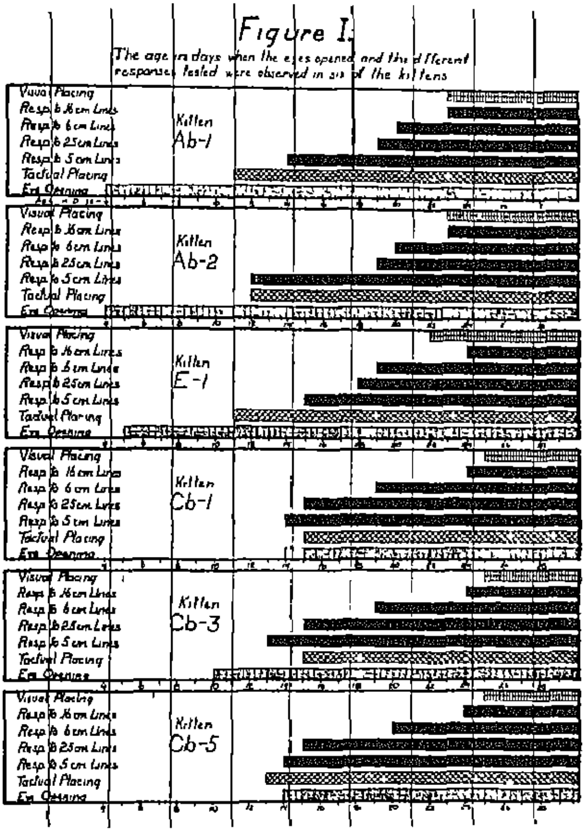


Figure I.

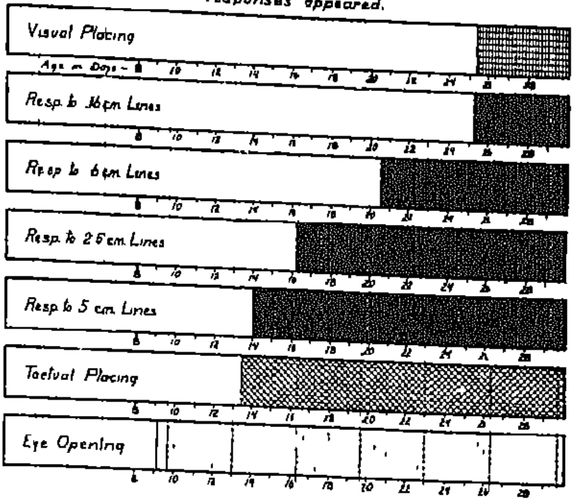
The age in days when the eyes opened and the different responses tested were observed in six of the kittens visual placing responses of the forelimbs is graphically shown in Figure 1, which presents a series of bar diagrams of the ages at which 6 representative kittens were observed to give evidence of the different responses studied. The length of a bar indicates the age at which the response listed at the left of the figure was first observed For the entire series of kittens used in this experiment, the rank difference correlation between the age of response to the 0 16 cm stripes and the age at which visual placing appeared is 0.93, which is a significant positive correlation for the number of cases presented A comparison of Tables 2 and 4 will show, however, that there were 5 kittens in all which developed visual placing reactions before responses to the 0.16 cm. stripes could be observed (kittens *Ca*-1, *Ca*-3, *D*-2, *D*-3, and *E*-1) Nevertheless, it must be noted that before any evidence of visual placing reactions could be seen, all of these kittens gave definite responses to the 0 6 cm stripes. The visual angle corresponding to these lines under the present conditions is 43 minutes of arc, an acuity of vision which is higher than that found by Lashley (21) in the adult albino rat by means of the discrimination technique. The present experimental results demonstrate, therefore, that a relatively high degree of visual acuity genetically precedes visually determined placing reactions

Figure 1 also summarizes graphically the age of the 6 different kittens at the time when all the various responses observed first occurred. It may be clearly seen that the age of eye opening (represented by the lowest bar in each of the six diagrams) shows a marked variation which does not appear to be correlated with either the age when visually determined head nystagmus was first observed or with the age when tactual placing reactions appeared This series of diagrams further demonstrates definitely that there is a gradual development of visual acuity in the kittens from the time that responses first appear to the stripes 5.0 cm. in width until the 0.16 cm stripes become effective in eliciting the oculocephalogyric reaction

Figure 2 represents the mean ages at which the various reactions listed toward the left appeared in all the animals used. The length of the bars indicates the mean age in days when the reactions were first observed in the whole series of kittens. As already seen previously in Figure 1, visual acuity gradually increases from the time that young kittens respond to the 5 0 cm stripes until these animals

Figure II

Mean ages in days at which the different responses appeared.

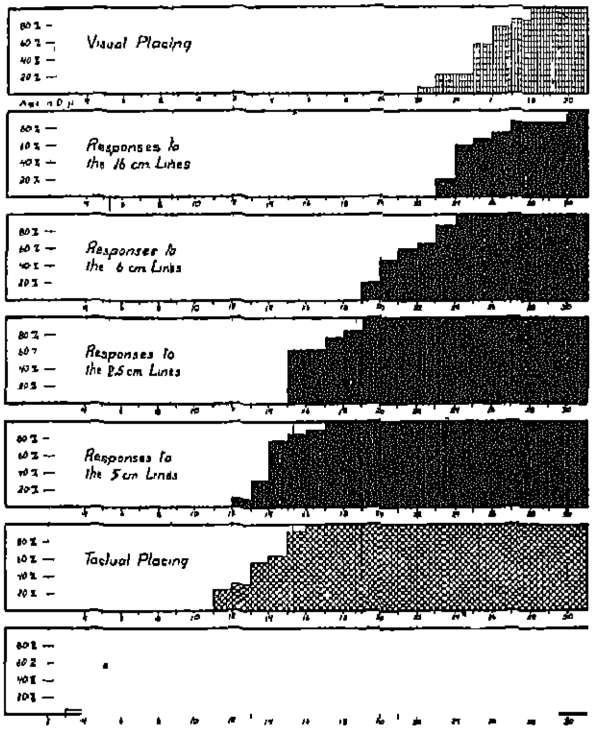


respond to stripes subtending a visual angle of 11 minutes of arc, and that the visual placing reactions do not appear until a relatively high degree of visual acuity is attained

Figure 3 presents the percentage of animals in which the different responses tested were observed on subsequent days The vertical component of the curve represents the percentage of all the animals studied in which the various reactions appeared at the different ages The horizontal component represents the age in days when the responses were first observed This figure serves to emphasize statements made in an earlier connection, namely, that the eyes opened at almost any time between the ages of 3 days and 15 days, and that the placing responses and the visual acuity increments show little variability in time of development among the different animals. This figure also shows clearly that there is a gradual increase in visual acuity with age Furthermore, little overlapping of ages seems to occur even when percentages of all animals used are thus represented. Finally, this figure serves to emphasize again the close correlation existing between the appearance of a relatively high degree of visual acuity and the development of visual placing reactions

Figure III

The percentage of animals in which the different responses tested were observed at subsequent ages in days



IV SUMMARY AND DISCUSSION

In order to ascertain the degree of visual acuity demonstrable at different age levels, 22 kittens were tested daily under controlled conditions from the time that the eyes first opened The visual acuity measurements were correlated with the age when the eyes opened and the age of onset of visual placing responses of the forelimbs.

The kittens used in this study opened their eyes at ages varying from 3 days to 15 days, with an average age of 9 days and a mean deviation of 3 days So far as could be ascertained in the present experiment, the age at which the eyes open seems to have no significant relation to the development of different visually determined reactions, to the time at which tactual placing responses first appear, and to the health or weight of the animal A negative rank difference correlation of 0 06 was found between the age of eye opening and the appearance of the oculocephalogyric response Notable in this connection is the fact that the mean deviation for the age when the eyes open is considerably larger than that for the genetic appearance of either the visually or tactually determined responses investigated

Certain important differences between the present data and the results and interpretations of prior investigators should be pointed out In general, it may be said that the results obtained place in question the conclusions of Soltmann (41), Held (14), Tilney and Casamajor (43), and Fish and Windle (8) regarding the close correlation between the age of eye opening and the onset of visually determined behavior For example, Tilney (43) states without qualification that,

> On the second day after birth, the optic nerve and tract of the kitten are entirely devoid of myelin sheaths, the eyes are then closed On the sixth day the optic tract is myelinized up to the superior colliculus, the pulvinar and the lateral geniculate body On the seventh day the animal opens its eyes There seems to be a difference in time of myelinization of perhaps 48 hours between the deposition of the myelin sheaths and the establishment of function

Although it is not known upon how many animals these and similar conclusions are based, it seems possible to say that the interpretations just described are open to considerable doubt In the present study the fact has been brought out clearly that the age of eye opening in the kitten is not restricted, as Tilney and Casamajor state, but can be found to vary from the ages of 3 to 15 days, or over a period of at least 12 days This variation has been shown to be unrelated to the age of onset of both oculocephalogyric reactions and visual placing reactions It can be concluded, therefore, that the neural growth changes associated with the onset of visual function in the kitten are entirely unrelated to the age of the opening of the eyes.

No satisfactory explanation has been given concerning the delay in development of visual function in kittens after the eyes have already opened As demonstrated by the experiments of Smith (39), the oculocephalogyric reaction does not necessarily involve cortical activity in the adult cat In further experiments, Smith (40) found that removal of the striate cortex in the adult cat does not abolish oculocephalogyric responses to a rotating striated pattern of lines subtending visual angles of at least 11 minutes of arc Tilney and Casamajor (43) described a subcortical mechanism for this reaction On the basis of these experiments it can be tentatively concluded that the development of visually determined oculocephalogyric reactions is not associated with growth and development of the visual cortex However, further experiments, involving operative procedures with young animals, are necessary in order to prove this contention

The data obtained in the present study of visual acuity has been summarized in Tables 4 and 5. There has been no previous study of the growth in visual acuity of young animals, and no explanation has been offered of this process of development which extends over a period of 9 to 18 days in the kitten (Table 5). These results show that the visual acuity of the kitten increased during this period from 360 minutes to 11 minutes of arc. It must be noted, however, that the data of the present experiment do not represent the maximal development of visual acuity in the cat (Smith, 38)

It was found that none of the animals studied showed any indication of visual placing reactions before they had given definite responses to the 0 6 cm stripes, and that 8 of the young kittens did not give visual placing reactions until they responded to the 0 16 cm stripes The mean age at which the responses to the 0.16 cm stripes appeared was the same as that for the first occurrence of visual placing responses, i e , at 25 days

A critical division in the neural control of visual placing reactions and the oculocephalogyric responses has been shown by Smith (39) to exist in the adult cat This division of function is somewhat similar to the genetic separation of these responses found in the present experiment In the adult cat, Smith has shown that removal of the visual cortex definitely abolishes visual placing reactions but leaves intact the oculocephalogyric responses, when both of these activities are tested under comparable stimulating conditions It seems possible to conclude from the experiments of Smith and from the observations made in the present study that the appearance of the visual placing reactions is correlated in time with the development of cortical connections of the visual system as well as with increased refinement of visual acuity, in so far as this may be mediated either cortically or subcortically.

Tactual placing reactions were observed to appear from 10 to 14 days earlier than similar visually determined placing reactions The mean age for the appearance of tactual placing responses was 13 days This is in accordance with the age at which Weed (46) and Langworthy (17) found decorticate rigidity in the forelimbs of kittens Bard and Brooks (4) localize placing reactions of the forelimbs of the cat in the gyrus proreus, the sigmoid gyri, the rostral portion of the longitudinal gyrus, and some fraction of the coronal gyrus Tilney and Casamajor (43) note that the gyrus sigmoideus is one of the two points of the cortex reached by myelinated fibers in the 14-day old kitten On the basis of these combined observations, the conclusion may be drawn that the period of time intervening between the primary development of the tactual placing reactions and the visual placing reactions represents the time of the formation of the visual connections with the cortical mechanisms of the placing responses. Since results and interpretations presented above point to the existence of a high degree of visual acuity preceding the appearance of visual placing reactions, it can be further stated that, prior to the establishment of optic connections with the motor cortex, a degree of visual acuity is already developed which permits the kitten to respond to lines 0 6 cm in width, located at a distance of 50 cm from the eyes.

V CONCLUSIONS

1 An apparatus was devised which could be used to present a rotating visual field to the kitten when the latter was placed on a stationary platform in the center of the rotating cylinder In addition, observations were made regarding the ages at which the kittens opened their eyes, and the ages at which tactual and visual placing reactions could first be elicited

2 By the use of the rotating visual field technique, the initial appearance of oculocephalogyric reactions was observed to occur in the different animals at ages varying from 12 to 17 days, with the mean at the age of 14 days

3 At the time of onset of these reactions the visual acuity of the cat in terms of the visual angle subtended by the stripes in the rotating pattern was 360 minutes of arc

4 From the date of onset of this visually controlled reaction, the visual acuity, measured in terms of visual angle, increased gradually in the following stages An acuity of 180 minutes of arc was demonstrated at ages varying from 15 to 19 days, with the mean at the age of 16 days Responses to stripes subtending a visual angle of 43 minutes were demonstrated at ages varying from 19 to 24 days, with the mean at the age of 21 days Responses to the smallest stripes, subtending a visual angle of 11 minutes, were observed in different animals at ages varying from 23 to 30 days, with the mean at the age of 25 days

5 Daily observations of the kittens showed that the eyes opened at the ages between 3 and 15 days, with the mean at the age of 9 days The data presented in this report show no relation between the opening of the eyes and the development of visual acuity in the cat A negative rank difference correlation of 0.06 was found between the age of eye opening and the age when head mystagmus was first observed

6 The development of visual acuity in the kitten, as measured in the present experiment, extended over a period of 9 to 18 days The course of development after an acuity of 11 minutes of arc has been reached remains to be determined in further experiments

7 Visual placing reactions of the forelimbs were observed to occur between the ages of 22 and 28 days, with the mean at the age of 25 days A positive correlation of 0 93 was found between the age when responses were observed to the 0 16 cm stripes and the age when visual placing responses could first be elicited. It may be concluded that this correlation shows a significant relationship, at least in time, between the appearance of visual placing reactions and the appearance of a relatively high degree of visual acuity, and that a relatively high degree of visual acuity precedes the appearance of visual placing reactions and other precise localizing responses of the limbs and trunk.

8 Tactual placing responses of the forelimbs appeared 10 to 14 days earlier than visual placing responses, and there is a possible correlation between the early appearance of tactual placing responses and an early appearance of visual placing responses It may be assumed

that the time intervening between the appearance of the tactual and the visual placing reactions represents the period of formation of functional connections between the visual and motor cortex, since evidence is available that both of these responses are mediated cortically (Bard and Brooks, 4, Bard and Orias, 5, Smith, 39) On the basis of such an assumption and the data already discussed, it may be further concluded that a high degree of visual acuity precedes the formation of functional associations involving the striate and motor cortex

REFERENCES

1 ALLEN, J The associative processes of the guinea pig *J Comp Neur* , 1904, **14**, 293-359

2 ANGULO Y GONZALES, A W Is myelinogeny an absolute index of behavioral capacity *J Comp Neur* , 1929, **48**, 459-464

3 AVERY, G F Responses of fetal guinea pigs prematurely delivered *Genet Psychol Monog* , 1928, **3**, 247-327

4. BARD, P , & BROOKS, C M Localized cortical control of some postural reactions in the cat and rat together with evidence that small cortical remnants may function normally *Proc Ass'n for Res. in Nerv & Ment Dis* , 1934, **13**, 107-157

5 BARD, P , & ORIAS, O Localized cortical management of visual and labyrinthine placing reactions *Amer J Physiol* , 1933, **105**, 2-3

6 BERGER, H Experimentell-anatomische Studien ueber die durch den Mangel optischer Reize veranlassten Entwickelungshemmungen im Occipitallappen des Hundes und der Katze *Arch. f Psychiat u Nervenkr* , 1900, **33**, 521-567

7 CARMICHAEL, L An experimental study in the prenatal guinea pig of the origin and development of reflexes and patterns of behavior in relation to the stimulation of specific receptor areas during the period of active fetal life *Genet Psychol Monog* , 1934, **16**, 337-491

8 FISH, M W., & WINDLE, W F The effect of rotatory stimulation on the movements of the head and eyes in newborn and young kittens *J Comp Neur* , 1932, **54**, 103-107

9 GARREY, W E A sight reflex shown by sticklebacks *Biol Bull* , 1904, **8**, 79

10 GOODMAN, L The effect of total absence of function upon the optic system of rabbits *Amer J Physiol* , 1932, **100**, 46-63

11 GRUNDFEST, H The sensibility of the sunfish Lepomis to monochromatic radiation of low intensities *J Gen Physiol* , 1931-32, **15**, 307-328

12 HECHT, S , & WALD, G The visual acuity and intensity discrimination of Drosophila *J Gen Physiol* , 1933-34, **17**, 517-547

13 HECHT, S , & WOLF, E The visual acuity of the honey bee *J Gen Physiol* , 1929, **12**, 727-760

14 Held, H Beitrag zur Kenntniss des Nervenmarks 3 Ueber experimentale Reifung des Nervenmarks *Arch f Anat u Physiol, anat Abt*, 1896, 222-229

15 Jacobsen, M M, Jacobsen, C F, & Yoshioka, J G Development of an infant chimpanzee during her first year *Comp Psychol Monog*, 1932, **9**, 1-94.

16 Jones, M C The development of early behavior patterns in young children. *J. Genet Psychol*, 1926, **33**, 537-585.

17 Langworthy, O R A physiological study of the reactions of young decerebrate animals *Amer J. Physiol*, 1924, **69**, 254-264

18 ——— Relation of onset of decerebrate rigidity to the time of myelinization of tracts in the brain-stem and spinal cord of young animals *Carnegie Instit. Wash*, *Contrib Embryol*, 1926, **17**, No. 89, 127-140

19 ——— The behavior of pouch-young opossums correlated with the myelination of tracts in the nervous system *J Comp Neur.*, 1928, **46**, 201-247

20 ——— A correlated study of the development of reflex activity in fetal and young kittens and the myelinization of tracts in the nervous system *Carnegie Instit. Wash.*, *Contrib Embryol*, 1929, **20**, No 114, 127-172

21. Lashley, K S Mechanism of vision III The comparative visual acuity of pigmented and albino rats *J Genet Psychol*, 1930, **37**, 481-484

22 Loeb, J Ueber die Summation heliotropischer und geotropischer Wirkungen bei den auf der Drehscheibe ausgeloesten compensatorischen Kopfbewegungen *Pflueg Arch*, 1907, **316**, 368-374

23 Lyon, E P On rheotropism I Rheotropism in fishes *Amer J Physiol*, 1904, **12**, 149-161

24 Magnus, R Koerperstellung Berlin J Springer, 1924 Pp xxi+740

25 McGinnis, J M Eye movements and optic nystagmus in early infancy. *Genet Psychol Monog*, 1930, **8**, 321-430

26 Michailow, S Zur Frage ueber die Erregbarkeit der motorischen Zentra in der Hirnrinde neugeborener Sauegetiere *Pflueg. Arch*, 1910, **133**, 45-70

27. ———. Die Entwickelung der Pupillen- und anderer Augenreflexe bei neugeborenen Sauegetieren *Pflueg Arch*, 1910, **133**, 71-81

28 Mills, W The nature and development of animal intelligence London. T Fisher Unwin, 1898 Pp xii+307

29 Mowrer, O H "Maturation" vs "learning" in the development of vestibular and optokinetic nystagmus *J Genet. Psychol*, 1936, **48**, 383-404

30 Muller, H R, & Weed, L H Notes on the falling reflex of cats *Amer J Physiol*, 1916, **40**, 373-379

31 Peiper, A Die Hirntaetigkeit des Saueglings Berlin J Springer, 1928 Pp 102

32 Pratt, K C. The neonate In *A handbook of child psychology*, ed by C Murchison Worcester, Mass Clark Univ Press, 1933 Pp 163-208

33 Radl, E Untersuchungen ueber den Phototropismus der Tiere Leipzig W Engelmann, 1903 Pp viii+188

34 Saelzle, K Untersuchungen an Libellenlarven ueber das Sehen bewegter Objekte *Zsch vergl Physiol*, 1932, **18**, 347-368

35 Schlieper, C Farbensinn der Tiere und optomotorische Reaktionen *Zsch vergl Physiol.*, 1927, **6**, 453-472

36 Shirley, M M. The first two years, a study of 25 babies Vol II Intellectual development Minneapolis Univ of Minn Press, 1933 Pp xvi+513

37 Small, W W Notes on the psychic development of the young white rat *Amer J. Psychol*, 1899, **11**, 80-100

38 Smith, K. U The visual acuity of the cat in relation to stimulus distance *J Genet Psychol*, 1936, **49**, 297-314

39 ——— The effects of removal of the striate cortex upon certain unlearned visually controlled reactions in the cat *J Genet Psychol*, 1937, **50**, 137-156.

40 ——— The effects of the removal of the striate cortex upon visual acuity in the cat as determined by oculocephalogyric responses *Psychol Bull*, 1936 (In press)

41 Soltmann, O Experimentelle Studien ueber die Functionen des Grosshirns der Neugeborenen *Jahrb f Kinderhlk*, 1876, **9**, N F, 106-148

42 Tilney, F Behavior in its relation to the development of the brain Part II Correlation between the development of the brain and behavior in the albino rat from embryonic states to maturity *Bull Neur Instit N Y*, 1933, **3**, 252-357

43 Tilney, F, & Casamajor, L Myelinogeny as applied to the study of behavior *Arch Neur & Psychiat.*, 1924, **12**, 1-66

44 Tinklepaugh, O L, & Hartman, C. G Behavior and maternal care of the newborn monkey. *J Genet Psychol*, 1932, **40**, 257-286

45 Turner, W D The development of perception. I Visual direction; the first eidoscopic orientations of the albino rat *J Genet Psychol*, 1935, **47**, 121-140

46 Weed, L H The reactions of kittens after decerebration *Amer J Physiol*, 1917, **43**, 131-157

47 Weed, L H, & Langworthy, O R Developmental study of excitatory areas in the cerebral cortex of the opossum *Amer J Physiol*, 1925, **72**, 8-24

48 ——— Physiological study of cortical motor areas in young kittens and in adult cats *Carnegie Instit Wash, Contrib to Embryol*, 1926, **17**, No 87, 89-106

49 Windle, W F Normal behavior reactions of kittens correlated with postnatal development of nerve fiber density in the spinal gray matter *J Comp Neur*, 1930, **50**, 479-503

50 Wolf, E Critical frequency of flicker as a function of intensity of illumination for the eye of the bee *J Gen. Physiol.*, 1933, **17**, 5-19

51 Wolf, E, & Zerrahm-Wolf, G The validity of Talbot's law for the eye of the honey bee *J. Gen Physiol*, 1935, **18**, 865-868

52 ——— Threshold intensity of illumination and flicker frequency for the eye of the sunfish *J Gen Physiol*, 1936, **19**, 495-502

University of Rochester
Rochester, New York

PERSONALITY RATINGS II. VALIDITY OF TEACHERS' RATINGS IN PREDICTING SCHOLASTIC ACHIEVEMENT*

Wesleyan University

T A Langlie

A previously published paper has shown that teachers' ratings are rather unreliable as shown by various criteria The present report is a study of the validity of the same ratings in predicting scholastic achievement at the University of Minnesota

Validity of tests is usually defined as the extent to which scores obtained represent true degrees of that which the test purports to measure. The validity of subjective estimates may be determined in like manner, namely by obtaining correlations between ratings and objective criteria For example, the validity of subjective estimates of intelligence of pupils may be determined by correlating such estimates with Intelligence Test scores, provided the Intelligence Test is valid

It will be impossible to obtain true validity coefficients in this study except possibly for Trait I "Ability to Learn," and Trait VIII, "Capacity for College Work," because we have no "objective" measures of the other traits included in our rating scale We will attempt, however, to obtain some indication of the validity of these ratings by the following methods

1 Determining the amount of agreement between individual raters
2 Comparing ratings on "Ability to Learn" with "College Ability Test" scores
3 Comparing ratings on all traits, and particularly ratings on "Capacity for College Work" with subsequent scholastic performance

Agreement between Individual Raters

The first criterion of validity which will be advanced is commonly spoken of as a criterion of reliability. A true criterion of reliability

---

*Received in the Editorial Office on December 17, 1936

would be in terms of self-consistency rather than agreement with other judges We shall employ this criterion as indicating the reliability of ratings, and also as indicating their validity, since personality traits such as "Initiative," "Dependability," "Industry," etc , are essentially social concepts. Since, then, these traits depend upon human judgment for their definitions, we will assume that, if different judges rate the same individuals in the same manner, they are rating them primarily in terms of the trait which is defined on the scale. That is, if a number of judges rate a group of individuals for "Dependability" and agree in their ratings, that group of individuals might be thought of as making up a scale showing various degrees of the trait in question In other words, the trait is defined in terms of social agreement

The extent to which this criterion is met by these ratings will be studied by the use of correlation technique

The data for this study comes from the ratings of Central High School for 1925 since there are a sufficient number of ratings of each individual in this group of ratings The reliability coefficients to be presented are not correlation coefficients between the ratings of one judge and the ratings of another judge, since no two teachers rated a sufficient number of the same students to give any reliabilities to such coefficients Instead the correlations obtained are between one series of ratings by several teachers with another series of ratings by other teachers For example, Teacher *A* may have rated students 1, 2, 3, 4, 5, Teacher *B* may have rated students 1, 5, 6, 7, 8, Teacher *C* may have rated students 2, 3, 8, 9, and 10, Teacher *D* may have rated students 4, 6, 7, 9, and 10 In the hypothetical example we have two ratings of each student, but no two teachers have rated more than two of the same students The correlations thus obtained probably come closer to true reliability coefficients than do the coefficients obtained by comparing only two judges who have rated the same people In a sense this method should give much the same result obtained statistically by Shen (6) Table 1 presents reliability coefficients as obtained from the 1925 data of Central High School using raw scores.

These reliability coefficients are far from satisfactory and indicate little agreement between any two ratings on the Graphic Rating Scale. Trait VIII, "Capacity for College Work," is rated the most reliably as represented by the coefficient of correlation of + 49 The

TABLE 1
RELIABILITY COEFFICIENTS, RAW SCORES MEN AND WOMEN, CENTRAL HIGH SCHOOL, 1925

Trait	Men r	Men P E	Women r	Women P E
I	31	05	.37	04
II	31	05	33	04
III	28	05	25	.04
IV	46	04	35	04
V	17	05	28	04
VIII	49	04	49	03
Total	40	04	41	03

reason for this is apparent,—namely that the high school records of the students are probably used as the basis for predicting future scholastic success. This may be back of the other correlations, also, but in lesser degree

A possible explanation of the low reliability of these ratings is that the teachers have not had the same experience with the students To the extent that such is the case this criterion is weak, since the raters should have similar contacts with the subjects, but some of these teachers will have had contacts with the students in English courses, others in mathematics, others in Social Science and other academic courses, and in a variety of extra-curricular activities as well. Such variation in opportunities to observe students would reduce agreement since the characteristics rated are probably quite specific in nature Thus one student may have an "aptitude" for English composition, but detests mathematics or chemistry, and hence shows to poorer advantage in courses involving mathematics His English teacher might rate him as "very superior" in "Ability to Learn" while his mathematics or chemistry teacher might rate him as "slow to learn." Other even more subjective traits might and probably would have even greater specificity

Kornhauser (3) has written as follows in regard to this point.

> Disagreements between ratings of the student may indicate that the student is different in different activities—not that instructors are wrong in their estimates What is called unreliability of ratings may really be in large measure a matter of the specificity of traits

It begins to appear that if Rugg's (5) requirements are necessary in order to obtain valid ratings, then valid ratings cannot be secured in the usual high school situation It is possible to obtain three ratings determined independently but not in substantial agreement Comparable and equivalent scales probably cannot be obtained by the Graphic method or by any other rating method since comparability is not present or absent in the scale, but is a function of the raters It is difficult to obtain teachers with adequate knowledge of the person rated, since no teacher observes the students in all situations Such lack of agreement as has been disclosed points definitely to the need for a greater recognition of the importance of demanding that teachers on every educational level should be "set" to observe and judge their pupils in addition to the time-honored function of imparting information

A personality rating on "Initiative," for example, indicates that some one is of the opinion that the person rated is characterized by a certain degree of this trait As defined on the Minnesota Rating Scale, "Initiative" should be judged by considering "resourcefulness and initiative in proposing new problems, applying principles, and making suggestions" Certainly a high school teacher is in a position to observe any manifestations of this trait in her pupils if that teacher is alert and capable of analyzing and interpreting human behavior Probably most teachers are not concerned with this method of approach in educational work, but would rather emphasize their subject matter, devoting all of their energies to making certain that their pupils master it.

Educational procedure is becoming more and more of a differential procedure, a conforming of subject matter and methods of instruction to the needs and capacities of individual students But recognition of personality differences is apparently lacking to a serious extent This lack is due, probably, to an inability on the part of teachers to observe definite behavior characteristics of their pupils in the absence of training and "set." This inability should be remedied by a more adequate training of teachers designed to make them "good judges of men" The need for such a program is urgent because guidance of pupils in terms of all aspects of their personalities is being increasingly stressed. As mentioned before, this need for personality diagnosis for guidance purposes cannot at the present time, nor probably in the immediate future, be met by objective personality tests As a result, teachers' ratings will be increasingly relied upon. Hence the need for some program which will improve their ratings.

Ratings on "Ability To Learn" Compared with College Ability Test Scores

Trait I, "Ability to Learn," approximates some of the more common definitions of "intelligence." Presumably the College Ability Tests were likewise designed to measure a similar function. For these reasons, we may investigate the validity of at least this one of the trait ratings by correlating "Ability to Learn" ratings with College Ability Test scores of the same pupils.

Data for this analysis were obtained from all three high schools, West, Central (1925 and 1926) and Humboldt. Table 2 presents these coefficients (Product-moment).

TABLE 2

Coefficients of Correlation between Teachers' Ratings on "Ability to Learn" and College Ability Test Scores

	West	Central '25	Central '26	Humboldt
		Single Ratings		
Boys	+.46 ± .06	+.40 ± .04	+.48 ± .04	+.54 ± .09
Girls	+.53 ± .06	+.47 ± .04	+.49 ± .03	+.49 ± .07

As shown by these figures there is only a fair degree of correlation between ratings on "Ability to Learn" and College Ability Test Scores.

Correlating the average of three teachers' ratings on "Ability to Learn" with test scores yields coefficients of +.53 and +.56 for men and women respectively (Central High School, 1925). It seems then, that teachers, rating individually, can judge the "Ability to Learn" of their students with some accuracy. When average ratings are used their accuracy is represented by a coefficient of correlation of about +.55. The reasons for such a relatively high validity of ratings in the face of previously demonstrated low reliability of these same ratings will be discussed in a succeeding section of this report.

Agreement of Ratings with Subsequent Scholastic Success

One Quarter Residence In the University. Single Estimates.

The final test of the value of teachers' ratings is to be found by comparing teachers' ratings with subsequent scholastic success. Achievement in a university is not a direct and unambiguous criterion of any of the traits included in this study, except Trait VIII, "Capacity for College Work," hence the relationships that exist between university scholarship and such traits as "Ability to Learn," "Initiative," etc, do not prove that the students have been rated accurately as to their possession of these traits, except in so far as each trait is a determinant of scholastic success But we have found that *there is a high degree of relationship between the several traits*, due, probably, to a "halo" effect We are not now concerned, then, with *the ability of teachers to judge these separate traits accurately*, but are rather concerned with the relationship which holds between their estimates of these traits and scholastic success On *a priori* grounds it is sound to expect that each of the six traits would correlate to a marked degree with success, *provided all other factors remained constant*

The subjects of this section of this report graduated from West High School in 1924 and from Central High School in 1925 and 1926 They entered the College of Science, Literature, and the Arts, University of Minnesota, in the fall immediately following their graduation from high school, and continued in residence for at least one quarter Most of them registered for 15 quarter credits, and, for the most part, their courses were similar, usually including Freshman English, a foreign language, and a science course As a whole, there was marked homogeneity as to courses taken Finer classification was not practicable, as the number of cases is none too large

Forty-six women, and 47 men who graduated from West High School, and were rated by at least one high school teacher, entered the University of Minnesota in September, 1924, and remained long enough to receive grades in their courses for the first quarter's work Twenty-seven women and 24 men who graduated from Central High School in 1925 and were rated by at least one teacher entered the University of Minnesota in September, 1925. All of the 192 students who are included in this section are native born and white

Scholastic achievement is determined by the method employed at the University for determining progress attained toward graduation, etc. Letter grades "*A*" to "*F*" are given in each course A grade of "*E*" is conditional and the final grade was then obtained at a later date in each case. In order to expedite the statistical treatment of these data the following numerical values were assigned the respective letter grades: *A*, 3 points; *B*, 2 points, *C*, 1 point, *D*, 0 points; and *F*, —1 point. These figures represent the quality of value of each credit obtained, hence a student who obtained a "B" in a five credit course would be given five credits and ten points (5x2). The final measure of average scholastic achievement is a ratio, obtained by dividing the total number of points earned by the total number of credits for which the person was registered. For example, a student who earned an *"A"*, a *"C"*, and a *"D"* in three five credit courses would receive an index of scholarship of (5x3) + (5x1) + (5x0) divided by 15, or 1.33. Thus a ratio of 1.00 indicates average or *"C"* work, a ratio of 2.00 indicates *"B"* work, an index of 3.00 *"A"* work, an index of .00, *"D"* work, and a ratio of —1.00, *"F"* work. These ratios will be called "Honor Point Ratios." Despite several objections to this method of indicating scholastic accomplishment, it is employed because the figures are readily available, and because this is the method used by the University Administration to determine the same facts about the achievement of students.

The students who entered in 1924 and 1925 were very similar to the student body of the University as a whole, as measured by scholastic achievement and College Ability Test scores. The 1926 entrants were superior to the usual run of students according to the same indices. These figures are presented in Table 3.

Despite a rather small sampling error, there being more "above average" than "below average" students among the subjects of this

TABLE 3

MEANS AND STANDARD DEVIATIONS OF THE VARIOUS GROUPS IN SCHOLARSHIP AND INTELLIGENCE, INCLUDING A CONTROL GROUP

Group		N	Honor Point Mean	Ratio S.D.	Intelligence Mean	S.D.
Control	Girls	310	.79	.87	48.81	27.60
	Boys	293	.61	.85	53.54	30.05
1924	Girls	46	.85	.83	53.00	26.10
	Boys	47	.67	.73	62.55	29.70
1925	Girls	27	1.18	.91	46.70	28.40
	Boys	24	.65	.78	59.46	28.10
1926	Girls	26	1.34	.78	72.50	25.70
	Boys	22	1.27	.74	69.50	27.30

study, the results obtained and to be presented would seem to be applicable, for the most part, to the rating by teachers of students in general as far as the University of Minnesota is concerned

The relationship between teachers' ratings and subsequent scholastic success are presented in the form of coefficients of correlation (Pearson product-moment formula) These coefficients are presented in Tables 4, 5, and 6

Despite the low reliability of these ratings, as shown in the first part of this study, there is a fairly high degree of validity, except

TABLE 4

Coefficients of Correlation Showing the Relationship between Teachers' Ratings and One Quarter of University Scholarship Central High School, 1925 96 Ratings on 27 Girls,* 86 Ratings on 24 Boys*

Traits	I	II	III	IV	V	VIII	Total
Girls	55 + 09	46 + 11	51 + 10	52 + 10	.25 + 13	68 + 07	.60 + 08
Boys	08 + 13	09 + 13	18 + 13	10 + 13	12 + 13	.22 + 13	13 + 13

*Probable errors of the coefficients of correlation have been computed conservatively on the basis of the number of pupils rated rather than on the basis of the number of ratings made

TABLE 5

Coefficients of Correlation Showing the Relationship between Teachers' Ratings and One Quarter of University Scholarship Central High School, 1926 71 Ratings on 26 Girls,* 60 Ratings on 22 Boys*

Traits	I	II	III	IV	V	VIII	Total
Girls	45 + 11	34 + 12	45 + 11	35 + .12	39 + 11	49 + 10	.48 + .10
Boys	43 + 12	34 + 13	33 + 13	.36 + 13	41 + 12	40 + 12	46 + 11

**Probable errors of the coefficients of correlation have been computed* conservatively on the basis of the number of pupils rated rather than on the basis of the number of ratings made

TABLE 6

Coefficients of Correlation Showing the Relationship between Teachers' Ratings and One Quarter of University Scholarship West High School 74 Ratings on 46 Girls,* 66 Ratings on 47 Boys*

Traits	I	II	III	IV	V	VIII	Total
Girls	.54 + 07	47 + 08	56 + 07	46 + 08	51 + 07	60 + 06	59 + 06
Boys	48 + 08	43 + 08	39 + 08	36 + 09	42 + 08	57 + 07	.55 + 07

*Probable errors of the coefficients of correlation have been computed conservatively on the basis of the number of pupils rated rather than on the basis of the number of ratings made for the 1925 group of men from Central High School The best intelligence tests yield coefficients of correlation of about + 60 with college scholarship These ratings correlate with scholarship to approximately the same degree On the face of it, there is a rather marked relationship between personality ratings and achievement, particularly between ratings on "Capacity for College Work" (Trait VIII) and University Scholarship

Ratings on particular traits correlate to a higher degree, or to the same extent, at least, as total ratings This suggests that *the ratings on "Capacity for College Work" are as valuable for prognostic use as ratings on all six traits* If further research does not contradict this result, then much time and energy would be saved by asking for ratings on a single general trait instead of on six or more specific traits, since the additional trait ratings add nothing to the value of the single general trait, as evidenced by the coefficients presented in the above tables This finding confirms the wisdom of present practice at the University of Minnesota in including this one general trait, "Capacity for College Work," on the credentials blank

Average of Three Estimates In order to increase the reliability of teachers' estimates the ratings of three teachers were averaged These average ratings were then correlated with University scholarship to determine whether such an increase in reliability would be accompanied by any marked increase in validity This analysis was confined to two classes from Central High School because there were too few students rated by as many as three teachers in the other high schools studied The results are presented in Table 7

These coefficients represent a fairly regular increase over the coefficients presented in Tables 4 and 5 They suggest that in pre-

TABLE 7

COEFFICIENTS OF CORRELATION BETWEEN AVERAGE RATINGS (3 RATERS) AND ONE QUARTER OF UNIVERSITY SCHOLARSHIP (HONOR POINT RATIO)

Traits	I	II	III	IV	V	VIII	Total
Central High School, Class of 1925 N= 25 girls 21 boys							
Girls	60 ± 09	52 ± 10	58 ± 09	55 ± 09	28 ± 12	74 ± 06	67 ± 07
Boys	09 ± 15	11 ± 15	20 ± 14	12 ± 15	15 ± 15	25 ± 14	15 ± 15
Central High School, Class of 1926 N= 23 girls 20 boys							
Girls	50 ± 11	39 ± 12	49 ± 11	39 ± 12	43 ± 12	55 ± 10	54 ± 10
Boys	48 ± 12	39 ± 13	36 ± 13	40 ± 13	46 ± 12	44 ± 12	52 ± 11

dicting scholastic performance, teachers' ratings should be averaged in order to increase their reliability, and hence their validity, though there are few statistically significant differences Some of the obtained high coefficients may be due to chance since there are very few cases involved, and yet the tendency is found in both classes and in all traits, suggesting that the tendency is definitely present This conclusion is in line with the findings of Rugg (5) in regard to ratings in the United States Army

One serious limitation of the demand that three independent ratings be averaged for each student is apparent from our own data, namely, —in practice it is difficult to obtain as many as three independent ratings on each subject In Industry it is almost impossible Even at the end of a four year high school course it would seem to be difficult although not impossible to find three teachers who feel sufficiently well acquainted with the average senior to warrant submission of a detailed rating

Long Time Residence In the University. The study of the ability of teachers to predict University scholarship so far presented utilized one quarter of University scholarship as a criterion We will now compare personality ratings with longer intervals of scholastic work

(*a*) *Two years of scholarship* Twenty-two boys and 26 girls who graduated from St Paul Central High School in 1926 and were rated by at least one teacher completed approximately two years of work at the University of Minnesota Honor Point Ratios, as described before, were used to represent their scholastic performance. These indices were then correlated with teachers' ratings on "Capacity for College Work" (Trait VIII) yielding a coefficient of correlation of + 32± 12 for the girls and + 61± 09 for the boys. The 22 boys were given a total of 60 ratings,—hence the correlation is between these 60 ratings and the Honor Point Ratio indices The 26 girls were given a total of 71 ratings, and these, likewise, are the measures which enter into the correlation scattergraph Probable errors which are given with the coefficients of correlation are computed conservatively *on the basis of 22 boys and 26 girls rather than* on the basis of the total number of ratings

Apparently, the prediction of capacity for college work of the boys is better when two years are considered than when one quarter is the criterion, but the prediction of success for the girls is not as high when two years are considered. This trend is unusual, since, in general, achievement of girls has been predictable to a greater degree than achievement of boys. Unfortunately, the small number of cases involved makes interpretation hazardous since we cannot be sure that the two coefficients would not fluctuate greatly in other similar samplings.

(*b*) *Three years of scholarship.* Three years after the 1925 classes from Central High School entered the University, many had cancelled out or were dropped for low scholarship. Some who cancelled their registration had honor point ratios above +1.00, which would indicate that they were doing satisfactory work at that time. On the other hand, some who were enrolled at the end of the third year had honor point ratios below +1.00, indicating that their scholarship was not satisfactory.

In the following table these students have been divided into five groups and their ratings on Trait VIII indicated. These groups represent (*a*) those who were enrolled at the end of three years and had honor point ratios of 1.00 or more, (*b*) those who were enrolled but had ratios of less than 1.00, (*c*) those who had cancelled even though their ratios were 1.00 or better, (*d*) those who had cancelled with ratios below 1.00, and (*e*) those who had been dropped by the University for low scholarship. These figures are presented in Table 8.

The right hand column of this table shows the per cent of ratings in each category which were equal to five scale steps or more. These measures are crude, but show a tendency which may be significant. Ninety-eight per cent of the ratings of girls who were enrolled at the end of three years and had honor point ratios of 1.00 or more were rated five or higher when they were high school seniors. Furthermore, all ratings of girls who had cancelled but had maintained a satisfactory scholastic standing (1.00 or higher) were above five. These two groups may be considered to be about equal in "Capacity for College Work" as shown by their grades in the University, since those included did do satisfactory work. Ratings show a similar relationship.

The two lowest categories are the fourth and fifth, "Cancelled with ratio of less than 1.00," and "Dropped for low scholarship." These two groups of girls were rated considerably lower than any of

TABLE 8

Showing the Relationship between Three Years' Scholarship and Ratings on "Capacity for College Work"

Central High School, 1925, N=27 Girls, 96 Ratings

Ratings, Trait VIII

	1	2	3	4	5	6	7	8	9	10	N	Median	Percent rated 5 or higher
Enrolled (1.00+)			1		5	5	12	9	7	6	45	8.00	98
Cancelled (1.00+)							3	2		3	8	8.5	100
Enrolled (1.00—)			2		1	3	3	1			10	6.7	80
Cancelled (1.00—)		2	5	3	9	7	6				32	5.7	69
Dropped (L.S.)				1							1	4.5	00
			Boys	N=24,	85	Ratings							
Enrolled (1.00+)			1	4	7	6	7	1			26	6.2	81
Cancelled (1,00+)				1	2	3	1	1			8	6.3	88
Enrolled (1.00—)			2	4	9	5	3				23	5.6	74
Cancelled (1.00—)			1	1	5	3	1				11	5.7	82
Dropped (L.S.)		1	3	1	8	5					18	5.5	72

the other groups although 22 of the 33 ratings assigned members of these two groups were five or higher.

The ratings of boys do not show as definite a tendency to correlate with achievement. This is in line with the low coefficients of correlation between ratings and first quarter scholarship as shown in Table 4, but is in marked contrast to the relationship between estimates and two years scholarship of the 1926 Central High School group. This discrepancy may be due to a sampling error, since the number of cases is small, but may also be due to lack of consistency of teachers' ratings.

(*c*) *Four years of scholarship.* At the end of four years some students have graduated, others have gone into professional schools, whereas others are still enrolled or have cancelled or have been dropped by the University. Table 9 presents the distributions of ratings of the

TABLE 9

SHOWING THE RELATIONSHIP BETWEEN FOUR YEARS' SCHOLARSHIP AND VARIOUS INDICES

West High School Girls N=46 Girls, 74 Ratings

Group	Ratings Trait VIII												
	1	2	3	4	5	6	7	8	9	10	N	Median	Percent rated 5 or higher
Degree						1	3	5	3	4	16	8.8	100
Enrolled (1.00+)					1	5	2	3			11	6.9	91
Cancelled (1.00+)				1			1	3	2		7	8.5	86
Enrolled (1.00—)				1		3	3	1			8	7.0	88
Cancelled (1.00—)			4	2	1	6	6	2	1		22	6.67	59
Dropped (L.S.)			2	2	1	3	2				10	6.00	50
	Boys N=47 Boys, 66 Ratings												
Degree					1	1	5	4		2	13	7.9	92
Enrolled (1.00+)					1	1	2	8		1	13	8.3	92
Cancelled (1.00+)						1	2				3	7.3	100
Enrolled (1.00—)		1			1	1	6	2			11	7.4	82
Cancelled (1.00—		2		1	5	3	5	1			17	6.2	53
Dropped (L.S.)				1	2	3	2	1			9	6.5	67

various categories as given in the preceding table with the addition of "obtained degree."

These figures show a rather decided tendency to show group differences, since girls in the "favorable" categories are rated 5 or more 100, 91, 88, and 86 per cent of the time whereas only 59 per cent of ratings of those girls who cancelled with poor scholarship and only 50 per cent of those who were dropped for low scholarship were rated 5 or more on "Capacity for College Work." Median ratings show the same tendency to discriminate.

The ratings of boys, while not quite as discriminating in terms of the per cent who were rated five or more, are also differentiated by their high school teachers according to their later performance to a fair degree Median ratings show this tendency more than the percentage figures

Combining the data from Tables 8 and 9 into one scattergraph shows a relationship as given in Table 9a

TABLE 9a

Rating

	1-2	3-4	5-6	7-8	9-10	N
Degree			3	17	9	29
Enrolled 1		6	31	44	14	95
Cancelled 1		2	6	13	5	26
Enrolled —1		10	23	19	0	52
Cancelled —1 or Dropped L S	5	27	61	26	1	120
N	5	45	124	119	29	322

On the basis of a crude analysis of these 322 judgments or ratings, it becomes apparent that a situation exists here which is similar to that which exists when scholarship is predicted by College Ability tests. That is, few individuals who are rated "low" do satisfactory work, whereas a considerable number who are rated "high" do poor work All students who were rated as low as 1 or 2 failed to do satisfactory work and were either dropped for low scholarship or cancelled their registration voluntarily No one who was rated as low as 3 or 4 obtained an Academic degree at the end of four years' work.

On the other hand, one student, with a rating of 9 cancelled with low scholarship and 45 judgments of 7 or 8 on "Capacity for College Work" were on students who had scholastic records below 1 00. Twenty-six of them were dropped for low scholarship or cancelled out and 19 were enrolled, but with low scholastic averages

These results suggest rather decidedly that teachers' ratings alone are fairly valuable for predicting scholastic performance over a long time period. In other words, such subjective estimates seem to possess a fairly high degree of validity

This conclusion is supported by correlations between ratings and one quarter scholarship, by correlations between ratings and two years scholarship, and by comparison of ratings with scholastic performance over three and four year periods

If such relationships obtain between ratings and scholarship for the total group of students, extreme cases should show the validity of teachers' ratings to a greater extent We will, accordingly, see what relationship holds between ratings and performance for extreme cases, and particularly for discrepancy cases

SENSITIVITY OF RATINGS IN DIFFERENTIATING DISCREPANCY CASES

(*a*) *Discrepancy between College Ability Test scores and University achievement* If teachers' ratings are valid measures of scholastic performance, extreme cases of high college ability and high scholastic attainment, low college ability, but high scholastic attainment, high college ability but low performance, and low ability and low performance, should be selected rather accurately Some of these are extreme cases of achieving according to tested ability whereas others are cases of marked discrepancy between ability and performance We wish to ascertain, then, how well teachers can predict scholastic achievement relative to tested ability. Table 10 contains data which will assist our analysis of this point

The records of 14 girls and 10 boys who graduated from Central High School in 1925 are presented in the upper part of the table and the records of 12 girls and 10 boys from the same high school in the senior class of 1926 are presented in the lower part of the table In such an analysis as is here undertaken the number of cases although admittedly small is sufficient to test the sensitivity of ratings in differentiating such marked discrepancy cases The anlysis is based on 46 cases in all, involving 127 ratings

We may interpret the table as follows Three girls in the *HI-HU* group, for example, were given eight different ratings which averaged 8 4 on "Ability to Learn," 7.6 on "Initiative," 7 8 on "Industry," 7.3 on "Scholastic Zeal," 7 6 on "Dependability," and 8 8 on "Capacity for College Work" The average scholastic performance of these girls is represented by an Honor Point Ratio of 2 22, and their average tested College Ability is represented by a percentile score of 95 These girls are a very select group, and relative to the other groups are rated in all traits as superior Their rating on "Capacity for College Work" is particularly high This group is clearly differentiated from the other three groups by average teachers' ratings

It is logical to expect that those who have low tested ability but achieve high grades will be relatively industrious, and that those who have high ability but do not do satisfactory work will be rela-

TABLE 10

Average Ratings of Students in the Four Quadrants of the Scattergraphs Showing Relationship between College Ability Test Scores and Honor Point Ratios

Group Selected		N	Ratings	I	II	III	IV	V	VIII	HPR	CAT	N	Ratings	I	II	III	IV	V	VIII	HPR	CAT
Central 1925	HI-HU	3	8	8 4	7 6	7 8	7 3	7 6	8 8	2 22	95	2	4	6 8	6 3	6 8	5 9	5 0	5 9	1 60	95
	LI-HU	1	3	6 3	3 0	4 0	4 0	4 7	5 7	1 67	24	0	0	—	—	—	—	—	—	—	—
	HI-LU	5	14	6 9	5 9	7 2	5 7	6 8	6.2	80	54	4	11	5 5	5 3	6 0	5 2	5 3	5 9	60	86
	LI-LI	5	13	4 6	4 3	5 5	4 3	4 9	4 5	—14	23	4	11	5 3	4 9	5 8	4 8	5 0	4 8	—35	14
Central 1926	HI-HU	4	13	8 1	8 0	8 4	7 8	7 9	8.0	2 43	91	4	11	8 2	7 7	7 8	8 1	8 3	8 1	2 32	95
	LI-HU	3	9	6 4	5 8	7 3	5 9	6 5	6 8	1 71	37	2	5	5 8	4 0	5 6	4 2	5 4	5 4	1 40	42
	HI-LU	2	5	6 6	7 0	6 2	6 2	5 8	6 4	74	91	1	3	7 3	7 3	8 3	8 7	8 3	8 3	24	93
	LI-LI	3	9	6 0	6 0	5 7	5 7	5 4	5 2	—11	26	3	8	5 3	5 3	4 8	4 5	4 8	4 6	46	28

Note HI-HU refers to the group who received high College Ability test scores and received high grades in the University
LI-HU refers to those who attained low College Ability test scores but received high grades in the University
HI-LU refers to the high College Ability and low scholarship group
LI-LI refers to the low College Ability and low scholarship group
Av HPR refers to the scholastic achievement
CAT refers to College Ability Test

tively lacking in industry This logic is not supported by teachers' ratings, however, since the one girl who was of low ability as tested but who achieved satisfactorily was given an average rating by three teachers of 6 3 on "Ability to Learn," of 5 7 on "Capacity for College Work," and only 4 0 on "Industry" and "Scholastic Zeal" On the other hand, five girls, who had relatively high ability as tested but did poor work in the University, were given 14 ratings which averaged 6 9 on "Ability to Learn," 6 2 on "Capacity for College Work," and 7 2 on "Industry"! Although given an average rating 6 higher than the *LI-HU* individual on "Ability to Learn," they were given an average rating of 3 2 higher on "Industry" Clearly, there is lack of differential rating sensitivity in this situation The trend which exists seems to follow tested ability rather than actual achievement, and this trend is found for all traits

Ratings of the 1925 group of boys seem to have little differential value. This is consistent with the low correlations obtained between ratings and scholarship for this group as given earlier in Table 5

Ratings of the 1926 group of girls reveal a similar tendency as was found in the ratings of the 1925 group of girls, with one marked exception The *HI-HU* group is differentiated from the other three groups in all traits, but the other three groups are rated at being quite homogeneous The ratings follow achievement rather than tested ability to a slight degree in some traits, particularly in Trait VIII, "Capacity for College Work," and "Industry" This time the *LI-HU* group is rated 7 3 on "Industry" as compared to 8 4 for the *HI-HU* group, 6 2 for the *HI-HU* group, and 5 7 for the *LI-LU* group Logically the *LI-HU* group should be even more industrious than the *HI-HU* group, since they achieve satisfactorily despite handicaps of low capacity Furthermore, it might be expected that the *LI-LU* group would be more industrious than the *HI-LU* group, but they are rated as lower

Ratings of the 1926 group of boys follow the tested ability of each group to a marked degree and do not differentiate according to achievement

These results suggest that very high ratings by teachers may be significant for use in predicting scholastic achievement, though they tend to follow tested ability more closely than actual achievement There is present, also, further evidence for a lack of analytical judgment in rating, i.e., teachers do not seem to judge according to specific merits and weaknesses of personality.

A PARTIAL ANALYSIS OF THE BASIS OF TEACHERS' RATINGS

In the preceding section we found that teachers' ratings tend to conform more closely to tested "Academic Ability" than to actual achievement. We found that presumably "lazy" students were rated about as high in "Industry" as hard working or normally industrious students. This finding suggests that the ratings are related more directly to the "examination value" of their students than to teachers' knowledge of the possession of various specific traits by their students.

In order to follow up this suggestion, coefficients of correlation were computed to determine the relationship between ratings and high school scholarship. Correlations between high school scholarship and university scholarship were also obtained. Partial coefficients of correlation between ratings and university scholarship, with high school scholarship held constant, were then computed. The correlations obtained between teachers' ratings and high school scholarship for those students who entered the University of Minnesota are presented in Tables 11, 12 and 13.

Apparently, according to the figures presented in the above tables, teachers' ratings correlate to a marked extent with high school scholarship. This relationship seems to be closer than the relationship between ratings and university scholarship, and gives some support to the suggestion which was mentioned earlier, namely that ratings may be based upon "examination value" instead of specific behavior characteristics.

Coefficients of correlation, zero order, between high school scholarship and university scholarship are:

West High School:	Girls +.65±.06	Boys +.65±.06
Central, 1925.	Girls +.74±.06	Boys +.32±.13
Central 1926	Girls +.69±.07	Boys +.50±.11

These correlations are about the same as those which usually obtain between scholarship in the University of Minnesota and high school records from large high schools such as West High school and Central High School. Variations which occur are probably due to the small number of cases in each group.

The next table, number 14, contains partial coefficients of correlation between teachers' ratings and scholastic achievement in the Uni-

TABLE 11

COEFFICIENTS OF CORRELATION SHOWING THE RELATIONSHIP BETWEEN TEACHERS' RATINGS AND HIGH SCHOOL SCHOLARSHIP

West High School 74 Ratings on 46 Girls,* 66 Ratings on 47 Boys *

Traits	I	II	III	IV	V	VIII	Totals
Girls	65±.06	66±.06	63±.06	63±.06	65±.06	70±.05	67±.06
Boys	62±.06	66±.06	54±.07	61±.06	58±.07	66±.06	63±.06

TABLE 12

COEFFICIENTS OF CORRELATION SHOWING THE RELATIONSHIP BETWEEN TEACHERS' RATINGS AND HIGH SCHOOL SCHOLARSHIP

Central High School, 1925 98 Ratings on 27 Girls * 90 Ratings on 24 Boys *

Traits	I	II	III	IV	V	VIII	Totals
Girls	62±.08	56±.09	54±.09	60±.09	49±.10	71±.06	67±.07
Boys	31±.12	26±.13	52±.10	40±.11	29±.12	47±.10	45±.11

TABLE 13

COEFFICIENTS OF CORRELATION SHOWING THE RELATIONSHIP BETWEEN TEACHERS' RATINGS AND HIGH SCHOOL SCHOLARSHIP

Central High School, 1926. 71 Ratings on 26 Girls * 60 Ratings on 22 Boys *

Traits	I	II	III	IV	V	VIII	Totals
Girls	58±.09	50±.10	61±.08	56±.09	50±.10	68±.07	66±.08
Boys	49±.11	45±.12	67±.08	.68±.08	66±.08	71±.07	69±.08

*Probable errors computed on basis of number of individuals, not on basis of number of ratings

TABLE 14

COEFFICIENTS OF CORRELATION BETWEEN RATINGS AND UNIVERSITY SCHOLARSHIP, WITH HIGH SCHOOL SCHOLARSHIP HELD CONSTANT

	I	II	III	IV	V	VIII	Total
			Girls				
West, 1924	20	07	25	09	15	27	27
Central, 1925	17	08	19	14	13	29	21
Central, 1926	09	00	05	—05	08	02	04
			Boys				
West, 1924	13	00	10	—05	10	28	29
Central, 1925	—02	01	02	—03	03	08	—02
Central, 1926	24	15	00	02	12	06	18

versity of Minnesota, when high school scholarship is held constant

By holding high school scholarship constant, the coefficients of correlation between teachers' ratings and university scholarship drop markedly These correlations suggest that teachers' ratings correlate with university scholarship only in so far as they correlate with high school records

The writer recognizes that there are pitfalls in the way of interpretation of partial coefficients of correlation Careful analysis on logical grounds of factors at work is necessary before any conclusions can be drawn as to cause and effect It would seem, however, that unless high school grades are determined by attitudes of "liking" or "disliking" on the part of the teachers, that teachers form their judgments in all traits included in this study on the basis of achievement in high school At least, judging on the basis of the practically zero coefficients of correlation which are obtained between ratings and scholarship with high school achievement partialled out, teachers' ratings add nothing significantly to high school records *for predicting university scholarship*

Further evidence that teachers' ratings and high school scholarship are more closely related than ratings and university scholarship is found by comparing ratings with location on the scattergraph representing the correlation between university scholarship and high school scholarship. This comparison will be similar to the analysis of ratings of discrepancy cases, except that we will not be concerned with marked agreement or disagreement between high school records and university scholarship The data for this analysis are contained in Table 15

A rather marked relationship is found here between high school achievement and ratings Those who had high standing in high school were rated high regardless of their later performance in the University Those who ranked low in high school were rated low, without differentiation according to their university performance This tendency is more pronounced here than it was in the comparison of discrepancy cases earlier.

Even though we cannot be sure of the nature of "cause and effect" in this situation we have evidence which leads us to believe that teachers' ratings have validity for use in predicting scholastic achievement only in so far as they represent achievement in high school It would seem that for use in guiding college students that teachers' estimates,

TABLE 15

RATINGS ON STUDENTS IN THE FOUR QUADRANTS OF THE SCATTERGRAPHS SHOWING THE RELATIONSHIP BETWEEN HIGH SCHOOL SCHOLARSHIP AND UNIVERSITY SCHOLARSHIP

Group Selected	N	Ratings	I	II	III	IV	V	VIII	Av H P R	Av H S P R
					Girls					
Central 1925										
HH-HU	5	13	8 6	7 9	8 2	7 8	8 1	8 9	2 33	91
LH-HU	2	5	5 4	5 0	5 4	5 8	5 4	5 8	1 07	23
HH-LU	2	4	6 3	5 5	7 3	5 3	5 8	6 0	94	80
LH-LU	4	11	4 5	4 2	5 2	4 1	5 6	4 2	— 20	16
Central 1926										
HH-HU	4	12	8 1	8 0	8 4	7 8	7 9	8 0	2 43	98
LH-HU	1	3	5 7	5 3	6 3	5 0	5 3	5 7	2 33	50
HH-LU	2	5	7 2	7 0	7 2	7 0	6 2	6 8	24	68
					Boys					
Central 1925										
HH-HU	4	11	6 6	6 3	6 7	5 8	4 9	6 1	1 60	79
LH-HU	0	0		None						
HH-LU	3	8	6 5	6 3	6 8	5 4	5 5	6 6	77	84
LH-LU	6	14	5 4	4 5	4 2	4 1	4 3	4 5	13	15
Central 1926										
HH-HU	2	6	8 7	8 3	8 3	8 5	8 0	8 7	2 57	95
LH-HU	1	3	5 0	3 0	4 3	3 0	4 0	4 7	1 33	11
HH-LU	2	5	6 4	6 4	7 6	7 4	7 0	7 0	28	71
LH-LU	3	9	5 7	5 6	4 7	4 1	4 6	4 6	22	18

Note HH-HU refers to the group who attained to high standing in high school and also at the university
LH-HU low high school scholarship, high university scholarship
HH-LU high high school scholarship, low university scholarship
LH-LU low scholarship in high school and in the university
Av H P R refers to the Honor Point Ratio
Av H S P R refers to High School Percentile Rank

when obtained under usual conditions and from large high schools, add no significant information to other more reliable measures of aptitude for college work In particular, they do not seem to contribute significantly to high school standing as a predictive device

SUMMARY

On the basis of the preceding analysis of teachers' subjective estimates of personality traits it appears that such estimates, although possessing a fair degree of validity for predicting university scholarship, do not add significantly to more objective measures, such as high school scholarship The evidence for this conclusion is as follows

1. There is a low degree of validity in terms of the personality traits being rated as evidenced by correlations between ratings by different teachers of the same students of from +.17 to +.40, with an average of +.35. One reason for such a low degree of agreement among raters in their estimates is probably due to lack of similar contacts with their students.

2. Ratings on "Ability to Learn" correlate to a fair degree (+.50) with College Ability Test scores, indicating a fair degree of validity of ratings on this trait. When averaged ratings are used their accuracy is represented by a coefficient of correlation of about +.55.

3. Personality ratings correlate with subsequent scholastic performance to a fair degree.

(*a*) Scholastic records for one quarter are predicted by ratings to an extent represented by coefficients of correlation of +.08 to +.68 with an average of +.38. "Capacity for College Work" ratings correlate with achievement to an extent of from +.22 to +.68, with an average of +.49. Averaging three ratings increases the correlation to some extent, though not significantly.

(*b*) Prediction of long time scholarship is about as accurate as prediction of one quarter's scholastic record.

4. Teachers' ratings do not differentiate those whose records on the College Ability Test are in disagreement with their university records. Ratings tend to follow test scores more closely than scholarship indices.

5. This tendency to be related to test records rather than to university scholastic records in ratings of discrepancy cases suggested that ratings are based on the "examination value" of the students. This suggestion was followed up by the partial correlation technique, and although conclusions as to cause and effect must be rendered with care, there is evidence that teachers' ratings do not add significantly to high school records in predicting university scholarship. This conclusion is based on partial correlations of from —.13 to +.29, with an average of +.10 between ratings and university scholarship, when high school records are held constant.

Ratings of students whose university and high school records disagree are related to high school records rather than university records.

DISCUSSION AND CONCLUSIONS

In this section attention will be given briefly to the general status of the problem of personality ratings. In this connection, certain implications arising out of our data and analyses will be discussed. Since discussion of specific findings has already been made in preceding sections, repetition at this point would seem to be unnecessary.

It is not necessary to review the history of testing to emphasize the unreliability and lack of validity of individual teacher's judgments of achievement, and yet their estimates of the personality and character traits of their pupils is commonly accepted by many college administrators and others. The situation is anomalous, since it is recognized by experts that teachers' judgments of achievement in particular subjects are liable to grave error when using essay type examinations, and yet they are expected to judge traits which are presumably infinitely more complex than achievement in spelling or arithmetic with sufficient accuracy to justify the use of those estimates in the guidance of college entrants.

According to the results of this study, teachers in large high schools, rating under conditions such as would be likely to exist in practice, do not judge students in regard to possession of personality traits with sufficient reliability to warrant the use of their estimates for placement and guidance of university entrants. Specific personality traits are not rated independently of each other, but each rating is affected by the general attitude of the teacher. This general attitude seems to be closely related to high school achievement. So much is this the case, that ratings correlate but negligibly with university scholarship when high school records are held constant.

Such results are in general agreement with the reports of other investigators. Kornhauser (4), for example, obtained coefficients of correlation between ratings of different teachers of +.38 and +.41. He obtained consistency coefficients for a few teachers who rated the same students twice. These coefficients averaged +.60, indicating better agreement with self than with others. He concludes,

> The agreement is not as uniformly close as might have been hoped for. There is evidence that instructors change their ratings of particular students considerably from one quarter to the next. To what extent this is reasonable and to be expected, we do not know. We are probably justified in interpreting it as due to insufficient knowledge about the people rated. But it may be that it represents, in part, actual changes in the behavior and attitudes of the students. In any case, this factor

is one of the important causes for the unsatisfactory final average ratings

The chief reasons for lack of reliability of ratings probably is traceable to the complexity and specificity of the traits rated, and to a confusion between *equivocance* and *equivalence* as pointed out by Johnson (1) and between the *jingle* and *jangle* fallacies, as pointed out by Kelley (2) That is, despite the use of similar names and the same descriptions, different teachers read them and understand them differently The remedy, obviously, would be to supply much more complete descriptions of the traits to be rated, and even then weights should probably be assigned to the elements which comprise the general trait Such a task is waiting accomplishment, and until it is accomplished judgments of personality traits will continue to be unsatisfactory If Johnson's article on "Some fallacies underlying the use of psychological 'tests' " (1) is sound, and the writer believes it to be so, then teachers' ratings, as at present obtained, are even more subject to attack on the same bases The marked lack of agreement among teachers when rating such traits as are usually included in rating scales is evidence for these conclusions No one would consider depending on objective tests with reliabilities as low as + 60 and yet teachers' ratings, which are no more reliable, are actually being used administratively in our schools and colleges

There is urgent need for personality analysis, however, and since few "objective" tests are available, teachers' ratings will be depended upon for some time to come But the solution of the problem will not be hastened by the indiscriminate use of rating scales. Careful training of teachers in habits of observation, analysis, and judgment is necessary

Before such training can result in satisfactory measures of personality traits, definition of those traits is imperative. In fact, analysis of behavior components, and formulation of definitions and specific descriptions would become the basis and furnish the content of a thorough plan for training teachers to observe, and enumerate significant phases of pupil behavior Teachers, psychologists, and personnel administrators should determine what specific behavior traits are to be included under each generic term. That is, they should determine what specific behavior traits are a manifestation of "Initiative," of "Industry," of "Dependability," etc In this way complete descriptions of each "trait" would be obtained and scales developed through research "Rating" by teachers would then merely consist of an indication of the presence or absence of each unit behavior trait in the same manner as they now determine the "rightness" or "wrongness" of each answer on a "newtype" examination After standardization of such scales, degrees of value for each unit trait could be determined by research, not by the judgment of each teacher when rating a particular student When this is done, ratings will no longer reflect the teachers' response to the student as much as the actual behavior of the student

The present situation with regard to the ratings of high school seniors turns out to be comparable to the situation found by Rugg in the rating of officers in the U S Army (5) At the end of his exhaustive study of army ratings, he wrote, "Study the data of this investigation as you will, one conclusion presses insistently—the complex traits of character must be measured objectively, not judged " Certainly public schools will find it administratively difficult, if not impossible, to secure three independent competent ratings on either a pupil or a teacher For diagnostic and training purposes helpful analyses can be made by question schemes But for the advancement of science, of education, and the better fitting of people for their life work and play we need to analyze character in great detail and to measure it objectively The measurement of dynamic personal and social traits stands today where the measurement of intelligence did 25 years ago—on a purely subjective basis Twenty years of laboratory measurement of specific functions preceded a decade of important synthetic work Since Binet's original synthetic scale of 1908 we have moved rapidly forward in the objective measurement of general intellectual ability, through the betterment of individual tests and the construction of group tests for measuring intelligence

But personality, character, temperament, disposition,—the fundamental complex attributes and their component dynamic traits—we have not even analyzed into their constituents, let alone made adequate tests for them" (Rugg, 5)

REFERENCES

1. JOHNSON, H M Some fallacies underlying the use of psychological tests. *Psychol Rev*, 1928, **35**, 328-337
2. KELLEY, T L Interpretation of educational measurements Yonkers World Book Co, 1927 Pp xiii+363
3. KORNHAUSER, A W A comparison of ratings on different traits *J Person Res*, 1926, **5**, 440 446
4. ———. A comparison of raters *J Person Res*, 1926, **5**, 328-334.
5. RUGG, H O Is the rating of human character practicable? *J Educ Psychol*, 1921, **12**, 425-438, 485-501, 1922, **13**, 30-40, 81-93
6. SHEN, E The reliability coefficient of personal ratings *J Educ Psychol*, 1925, **16**, 232-326.

Wesleyan University
Middletown, Connecticut

THE MENTAL DEVELOPMENT OF CHILDREN FROM UNDERPRIVILEGED HOMES*[1]

State University of Iowa

HAROLD M. SKEELS AND EVA A. FILLMORE

PROBLEM

Familiarity with the pattern of *IQ*'s of children in the same family when admitted to an orphanage suggested the present study. Scanning of family records made it seem pertinent to investigate whether there was really an unusual, consistent downward trend of *IQ*'s among the older members of a family, whether the length of stay in the type of home represented might be related to mental status as measured by the Stanford-Binet examination.

SUBJECTS

As part of the research program of the Iowa Child Welfare Research Station in cooperation with the Iowa State Board of Control, children are given mental examinations as soon as possible after entrance to the state orphanages. The examiners are staff members who have had long experience in mental testing and dealing with children in varied situations. For this study records were used from one orphanage since its children provided an age range satisfactorily wide for the purpose. There were 132 families, each having two or more children who had been tested within a short time after entrance. The large number in residence at its beginning were lost to this study, since only those tested promptly after entrance could be considered. As a rule the tests were made within a few days or at the latest a few weeks after entrance. These families provided 407 children ranging in age from one to 14 years. The largest family had eight children with the mean number per family, three.

---

*Received in the Editorial Office on December 18, 1936.

[1] The writers wish to express their appreciation to the Iowa Board of Control of State Institutions which at all times cooperated in making possible research studies of state wards, to Mr. Syl McCauley, Superintendent of the Iowa Soldiers' Orphans' Home, at which institution the children in this study were examined, and to Dr. George D. Stoddard, Director of the Iowa Child Welfare Research Station, who has made possible this study.

FAMILY BACKGROUND

A necessary part of this research was the investigation of the background in which these children had been reared. Whatever might be shown about their intelligence must be interpreted in relation to the home situation. Material about the family and home was found in the family histories sometimes submitted as part of the official procedure of entering a child in the institution. Such records contained a certain amount of factual data about the parents as well as descriptive accounts of an impressionistic nature. These records were in some cases very fragmentary, but there were enough facts available for the group to give a picture considered adequate for the purposes of the study.

Status of Parents

(*a*) *Occupations*. For 71 families the occupation of the father was given in the history. These were rated according to the classification of occupations of employed males in the United States from the report of the *Fourteenth Census of the United States*, Volume 4, 1920 [cited by Goodenough (2, Appendix, pp. 501-502)], as shown in Table 1.

TABLE 1

Group	Number	Total	Per cent
1	0		
2	0		
		0	
3	6		
4	2		
5	15		
		23	32.4
6	12		
7	36		
		48	67.6

Judged by this standard, 67.6 per cent of the fathers were employed as day laborers, slightly skilled workmen, or in occupations requiring little training or ability. The remainder were in semiskilled or skilled occupations, clerical positions, business, and farming. None were in the ranks of professional occupations. The 67.6 per cent in the lower ranking occupations may be contrasted with the 32.2 per cent of such workers in the nation as a whole. On the basis of occupational level, it might be assumed that constructive opportunities for the children had probably been limited. Such an assumption appears valid when facts enumerated later are also considered. Although the occupational status was given for only a part of the group, these families seemed to have enough characteristics in common with the others to make them a typical sample of the 132 families.

(*b*) *Education.* The educational status of the parents is shown in Table 2.

TABLE 2

School Grade Reached	Both Parents Mothers	Both Parents Fathers	One Parent Mothers
2	2	0	2
3	1	2	3
4	1	3	1
5	2	0	2
6	4	3	7
7	3	4	3
8	15	16	20
9	3	0	4
10	5	5	6
11	1	0	1
12	1	4	2
Some college	0	1	0
Total	38	38	51
Mean	7.5	8.0	7.5
Median	8.0	8.0	8.0

For 38 families records of both parents were given. The grade range of the mothers was two to 12, with a mean grade of 7.5. The fathers' range was three to 12, with one having an unstated amount of college training; their mean was grade eight. The median for each group of 38 parents was grade eight.

When the record for mothers as a group were considered there were 51. The mean was again grade 7.5 and the median again grade eight.

(*c*) *Residence in Institutions.* In these 132 families, of the 111 mothers living at the time the children entered the orphanage, 18 or 16.2 per cent were residents of state or county institutions. Eight were in insane hospitals, five in the reformatory, two in homes for feeble-minded, and three in county homes for the indigent. Of the 119 living fathers, 19 or 17.1 per cent were residing in institutions

Nine were in the penitentiary, four in jails, four in insane hospitals, and one in the reformatory

In addition to the foregoing group of parents, there were five mothers and 18 fathers who at some time had been in institutions Thirteen had been in jail, four in the penitentiary, four in insane hospitals, and two in the reformatory

Desertion of Family Eleven families had been deserted by one of the parents In two families both parents had left

Family Situation Although based on evidence for only a part of the families, it is considered that the findings with respect to the status of parents are typical of the whole group The descriptions of the homes, the characteristics of the parents, and the estimates of their capacities give a wider range of evidence than the numerical data, and it all creates a picture of sameness It is felt that had there been more records of occupations and education they would not have changed the levels as found The group appears a homogeneous one in which parents were largely in the lower occupational levels, with the average education being the eighth grade While poor economic status and lack of educational opportunities need not necessarily be considered damaging, there were in these families other factors that could not be compensated for The apparent low mentality of many parents, the poor condition of the homes, lack of parental responsibility, and the presence of insanity and law-breaking activities created circumstances of underprivilege to an extreme degree. We conclude then that we have a group of children reared in homes quite uniformly of a type that could not provide more than a minimum of stimulation toward mental development.

MENTAL STATUS OF CHILDREN

IQ Distribution for Entire Group With a mean *IQ* of 88 5, this group of 407 children, one to 14 years of age, is seen to represent a lower level of intelligence than prevails in the population at large Taking the orphanage entrants, five to 14 years of age, in comparison with an unselected group of the same age range (3), the extent of the lower *IQ*'s is shown (Figure 1).

IQ Distribution by Age When considered by age, it is seen by the tabulation in Table 3 that the groups which reach or exceed the mean for the whole, 88 5, are under 10 years of age

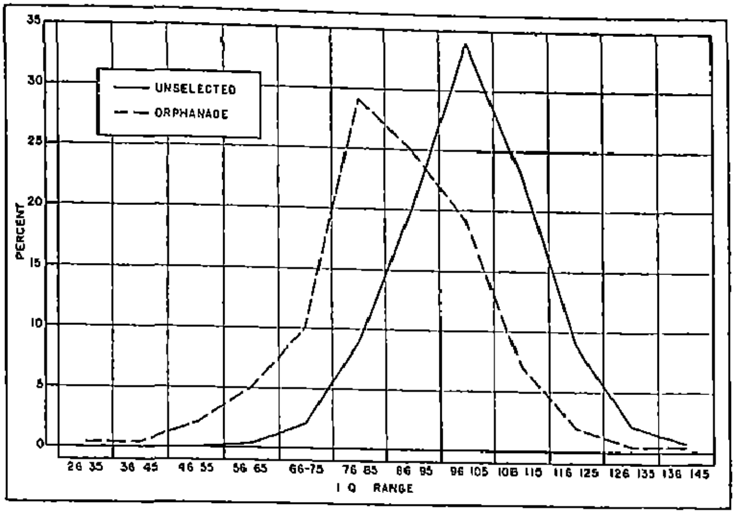


FIGURE 1
Distribution of *IQ*'s of 330 Orphanage Entrants and 905 Unselected Children, Ages Five to 14 Years
[Fig 2, p 66 (3)]

TABLE 3

Chronological age, years	Number	Mean IQ
1	8	108 4
2	27	93 3
3	22	87 7
4	20	92 6
5	42	89 7
6	24	98 0
7	37	92 2
8	35	90 7
9	51	88 7
10	34	85 4
11	37	83 3
12	35	81 6
13	17	79 6
14	18	79 9

There is no steady trend in *IQ* among the younger age groups, but beginning with year eight the mean *IQ* is lower than that for the preceding age groups combined. The downward trend then continues.

The distinction between the two age groups, below and above eight years, is seen when they are compared with an unselected group (3) (Figure 2). Since the unselected group begins with year five, for this comparison the orphanage group is begun at this point likewise. While both orphanage groups show a decided swing downward from the normal pattern, with higher percentages of lower *IQ*'s, the deviation is much more marked for the older group.

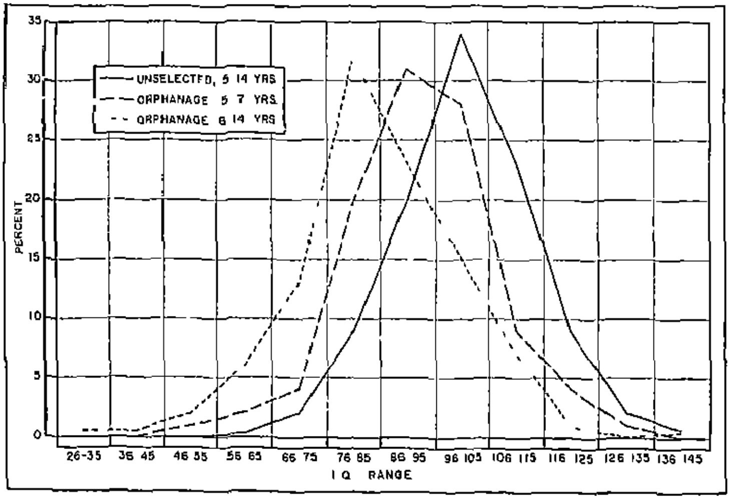


FIGURE 2
DISTRIBUTION OF *IQ*'S BY AGE OF ORPHANAGE AND UNSELECTED CHILDREN
[Fig 2, p 66 (3)]

Since a slight decrease in *IQ* for successive age groups has quite commonly been found, it remains to be seen how the decrease for the orphanage group compares with that for unselected children. For comparison the *IQ*'s by age of Terman's 905 unselected children are used (1). While they show a slight fluctuation below their mean of 102.0 beginning with year 12, the orphanage group shows a greater one below their mean of 87.5, beginning with year 10 (Figure 3).

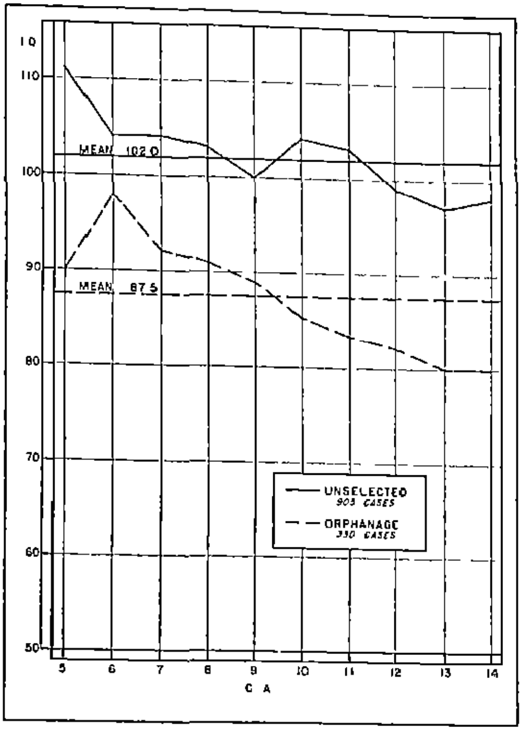


FIGURE 3
MEAN *IQ* BY AGE OF 330 ORPHANAGE ENTRANTS AND 905 UNSELECTED CHILDREN
[From table in 27th Yearbook (1)]

IQ and Length of Stay in Underprivileged Home. Since the age of the children as used in the preceding summaries was taken as of the time of entrance to the orphanage, age therefore corresponds to the length of time spent in the child's own home Homogeneous as this group has been shown to be in respect to family background, it is considered that the findings for age groups are one indication of the relation of *IQ* to length of time in the home. While it is known that the younger children in these homes have not been in stimulating surroundings, the effect of this is a matter of time as these results suggest Although there are fluctuations by age among the mean *IQ*'s of the younger children, on the whole they maintain a higher level than do the older ones However, after seven years in homes of a decidedly poor character, the mean level of intelligence drops with each successive age group With the differences between each age so slight as to be insignificant, the emphasis is, of course, put upon the trend as a whole Although lengthening the child's stay in the home by one year apparently makes no difference, a comparison of age groups two years or more apart suggests that among children in the upper age levels it may be possible to determine a relation between the increment of stay and the extent of *IQ* decrease Gross comparisons of the means of various age groups suggest that when children have been in these homes for 11 years, a mental level has been reached from which there is little deviation (Table 4)

TABLE 4
Comparison of Mean *IQ*'s of Various Age Groups

Age Groups, Years	Mean IQ Younger	Mean IQ Older	Difference
	Two Years Difference		
8 and 10	90.7	85.4	5.3
9 and 11	88.7	83.3	5.4
10 and 12	85.4	81.6	3.8
11 and 13	83.3	79.6	3.7
12 and 14	81.6	79.9	1.7
	Three Years Difference		
8 and 11	90.7	83.3	7.4
9 and 12	88.7	81.6	7.1
10 and 13	85.4	79.6	5.8
11 and 14	83.3	79.9	3.4
	Four Years Difference		
8 and 12	90.7	81.6	9.1
9 and 13	88.7	79.6	9.1
10 and 14	85.4	79.9	5.5

Mental Status of Siblings as Related to Length of Stay in Home

Method of Comparison To investigate further the relationship between length of time in the child's own home and *IQ*, children in the same families were paired with each other Each child was compared with every other one in his family, record being made of difference in age or length of time in the home and the *IQ*'s of the older and younger members of each pair Since the family backgrounds were homogeneous, all the data provided by these pairs of siblings were used in the various classifications. This method resulted in the older or first and second members of the family appearing in more pairs than the younger ones. This would be objectionable if the older ones were largely from the upper age groups in which the lower *IQ*'s prevail, in which case an overweighting downward would occur in the *IQ*'s of the older members of the pairs. However, this did not happen in this group since the first and second children were of all ages from two to 14 years. Moreover, the distribution on the basis of age differences (various duration of stay) provided a further means of offsetting unfair concentration.

IQ AND VARIOUS LENGTHS OF STAY IN HOME FOR CHILDREN OF ALL AGES

As previously stated, age differences between siblings were equivalent to the time spent in their own homes. For the pairs these differences showed the length of time which the older members of a family had been in the home beyond that of the younger members. For each time difference there was also the *IQ* of the older and younger child for whom the difference had been obtained. On the basis of time differences in units of years, mean *IQ*'s were computed for all older members and younger members of the pairs in each time category.

The results of the comparisons of *IQ*'s (Table 5) indicate that a difference in length of stay in the home of one year is not joined with a significant difference in *IQ*. When the length of stay is two years longer for the older members, then the *IQ*'s are lower to an extent very near to complete reliability. However, when the difference is extended to three years, it is then reliably established that the children who have been in the home that much longer than their siblings actually have lower *IQ*'s. With wider intervals of years' stay this difference between the older and younger children is sustained, although with not quite as high a certainty of significance, perhaps due to the fewer individuals represented.

IQ AND LENGTH OF STAY IN THE HOME FOR SELECTED AGE GROUPS

Further analysis of the relation of *IQ* to the length of time children stayed in their homes was made by comparing groups of siblings of specific ages. Since small differences in *IQ* were found for children

TABLE 5

DIFFERENCES BETWEEN MEAN IQ'S OF CHILDREN WHOSE STAY IN THE HOME IS SEPARATED BY VARIOUS LENGTHS OF TIME

Length of time in Home, Year	Pairs	Mean IQ Older Pairs	Mean IQ Younger Pairs	Standard Deviation of Distribution Older Pairs	Standard Deviation of Distribution Younger Pairs	Standard Deviation of Mean Older Pairs	Standard Deviation of Mean Younger Pairs	Standard Deviation of Difference	Difference in Mean of Older and Younger Pairs	Ratio of Difference to Standard Deviation of Difference
1*	66	84.15	86.82	12.50	12.90	1.54	1.59	2.21	2.67	1.21
2	140	86.26	90.10	15.40	14.70	1.30	1.24	1.80	3.84	2.13
3	80	85.69	93.01	12.65	17.05	1.42	1.91	2.38	7.32	3.10
4	93	82.51	88.74	13.70	12.80	1.42	1.33	1.95	6.23	3.19
5	66	84.83	92.35	14.75	16.55	1.82	2.04	2.73	7.52	2.75
6	34	85.08	93.82	13.88	18.50	2.38	3.17	3.96	8.74	2.21
7	34	81.55	89.82	13.00	14.95	2.23	2.56	3.39	8.27	2.44
8	24	82.83	94.62	16.00	18.20	3.26	3.71	4.94	11.79	2.39
9	16	81.37	96.12	15.35	16.00	3.84	4.00	5.54	14.75	2.66

*In the category one year, children remaining in the home from six months, zero days to one year, five months, twenty-nine days are included

TABLE 6

'DIFFERENCES BETWEEN MEAN IQ'S OF SIBLINGS OF SELECTED AGES

Age Groups, Years	Number	Mean Chronological Age	Mean IQ	Standard Deviation of Distribution	Standard Deviation of Mean	Standard Deviation of Difference	Ratio of Difference to Standard Deviation of Difference
1 to 7	99	5.0	96.0	16.5	1.67		
8 and 9	63	8.5	89.9	13.4	1.70	2.38	2.56
1 to 7	74	5.3	93.1	15.4	1.79		
10 and 11	51	10.5	84.2	12.9	1.81	2.55	3.49
1 to 7	45	5.3	92.6	15.0	2.23		
12 and 13	26	12.4	78.9	12.7	2.48	3.33	4.09

under eight years of age, it was considered of interest to determine how they might differ from siblings of the following ages: eight and nine years, 10 and 11 years, and 12 and 13 years respectively The results for these selected groups follow the same tendencies noted for the children as a whole (Table 6). While the groups below eight years shift, inclusion in them depending upon having siblings of the prescribed higher age range, the mean *IQ* does not vary much Those in the same families who have been longer in the home show a decreasing mean *IQ* The differences in *IQ* between younger and older children are seen to increase with the lengthening of the age difference or the time which the older group has been in the home in excess of the younger group Time differences of 3 5 years, 5 2 years, and 7 1 years likewise correspond to increasingly significant differences between younger and older children in respect to mean *IQ* Thus it may be inferred that in the type of home represented older children are definitely and increasingly inferior to those siblings under eight years of age

Summary

A group of 407 orphanage entrants belonging to family groups have been studied with reference to home background and the development of intelligence in relation to the type of homes represented The findings are as follows

1 These children came from uniformly poor backgrounds, as indicated by low economic status, limited education of parents, and conflict with the law

2 The level of intelligence of the children, one to 14 years of age, is shown by a mean *IQ* of 88 5

3. Children under eight years of age are of a higher mental level than those older

4. The *IQ*'s of the older children decrease with age to a greater extent than is found for unselected children, suggesting the retarding effect of poor homes on mental development

5. For siblings of all ages a difference of three years' stay in the home corresponds to a significant difference between the *IQ*'s of pairs separated by that difference

6 As the length of the stay in the home increases, the mean *IQ*'s of groups likewise decrease in a significant degree

REFERENCES

1. FREEMAN, F. N., HOLZINGER, K. J., MITCHELL, B. C., et al. The influence of environment on the intelligence, school achievement, and conduct of foster children. *27th Yearbook, Nat. Soc. Stud. Educ.*, 1928, 103-217.
2. GOODENOUGH, FLORENCE L., & ANDERSON, J. E. Experimental child study. New York: Century, 1931. Pp. xii+546.
3. TERMAN, L. M. The measurement of intelligence. Boston: Houghton Mifflin, 1916. Pp. xviii+362.

Iowa Child Welfare Research Station
State University of Iowa
Iowa City, Iowa

THE FREE WORD ASSOCIATION OF ELEMENTARY SCHOOL CHILDREN*[1]

North Carolina State College

WILLIAM MCGEHEE

I REACTION TIMES

The present investigation has developed out of a search for an adequate word list to be used in examining elementary school children by the free word association method An examination of the literature concerning the free word association method revealed that no adequate investigation has been reported which deals with the reaction times of children to any stimulus list The studies of Woodrow and Lowell (13), the Rosanoffs (8), and Wheat (11) omitted any comprehensive investigation of reaction times The early study of Rusk (9) presents data from only 22 children ranging in age from seven and a half to fourteen and a half years The psychological clinician, therefore, who uses the free association technique in examining children has no basis of fact to which to pin his conclusions regarding the reaction times of children.

The purpose of this investigation, therefore, is to study the reaction times of children of elementary school age in order to determine if either sex or age differences are revealed by differences in reaction times Secondarily, the present investigator will present any etiological factors of reaction time which are found in the course of investigating the major problem.

It would not be remiss, then, to sum up briefly the etiological factors which earlier investigators have discovered in connection with reaction times of adults as well as children Jung (4) unconditionally attributes long reaction time to emotional inhibition. Hull and Lugoff (3), Cason (1), White and Powell (12) point to the effective tone of the stimulus word as one of the factors in long reaction time Cattell (2) found that reaction time differed according to the type of response made but also that the familiarity of the subject with the stimulus word, his previous training and the frequency of usage of

---

*Received in the Editorial Office on December 28, 1936

[1] The author acknowledges the valuable assistance of Dr Paul L. Boynton, George Peabody College, in preparing this study

the stimulus words were strong etiological factors. Otis (6) found the main factor in the case of children's differences in reaction times is mental ability. Miles and Terman (5), in their summary of the literature dealing with sex differences as revealed by the free association technique, found that women reacted slower than did men. Rusk (9) reports a decrease in length of reaction times with an increase in the chronological ages. Wells (10) perhaps presents the best point of view in his summary of the etiological factors involved in differences in reaction time. He says,

> The free association time is exactly what the name implies, a measure of association freedom. It involves a freedom of presentation, of selection, of expression. We obtain uniformly short association times in those individuals in whom all these processes are comparatively free. Where longer times are found, obstruction of some sort is indicated. We have traced some of the characteristics of this obstruction, of its range of variability.
>
> This is as far as the unsupported association time will carry one. It remains then to become acquainted with the nature of the obstruction. Where it is found, it must not be incontinently set down to emotional factors, but the roles of suppression, indecision, and the like must be carefully compared and weighed by the best available means. Then only do differences in association time assume their complete significance for the test, which is the most searching of all psychological experiments, and gives more coercively than all others, a paradigma of the subjects' mental habit (10, 23).

The word list used in this investigation consists of 55 stimulus words selected by the members of an advanced class in child psychology at George Peabody College as words which not only are familiar to children and common to their experience, but which, still further, are sufficiently personal to avoid the transitory, almost hit-or-miss types of reactions which are so dominant in the responses of many children to certain previously used word lists. The word list is as follows:

1. money
2. afraid
3. cold
4. mad
5. dance
6. angel
7. doctor
8. police
9. bed
10. dog
11. prayer
12. boy
13. doll
14. cheat
15. ghost
16. prison
17. dream
18. flag
19. girl
20. dark
21. papa
22. friend
23. fun
24. fuss
25. sick
26. run
27. frog

28 milk
29 game
30 teacher
31 sing
32 young
33 love
34 star
35 whip
36 circus
37. coward
38 mother
39. tunnel
40. slap
41 pretty
42 help
43. park
44. hungry
45 Sunday
46 go
47 happy
48 mamma
49 school
50 laugh
51 stealing
52 water
53 home
54 book
55 baby

The words were presented orally to the subject and records of both his response and reaction time were made by the examiner. As near as possible test conditions were kept the same for each subject

The reaction times for failures to respond were not recorded but the omission of these times will not affect materially the reliability of any conclusions in view of the fact that failures occurred only 694 times in 22,000 responses

The reaction times were measured with a stop watch graded to fifths of a second The watch was started when the stimulus word was pronounced by the examiner and stopped when the subject made his initial response The subjects were instructed to reply with only one word and to reply as quickly as possible The following instructions were read to him.

> I am going to say some words to you, one at a time When I say a word I want you to answer just as quickly as you can with the first word that my word makes you think of Just say one word, and say it as quickly as you think of it For example, if I said the word "candy" you might say "good" or "eat" or "buy" Any one of these words would be right, because I only want to know what you think Answer with just one word, and say it as quickly as you think of it What would you say if I said "grass"?

The subjects for this experiment were 400 children of ages 7-10 residing in Nashville, Tennessee, and its immediate vicinity and in Paris, Tennessee For the sake of making comparisons more reliable and consistent, 50 children in each sex-age level were tested. These children were all either in public or private schools or in public or private institutions for the care of orphan children

Table 1 shows the median age of the various sex-age groups.

It will be noticed that in none of the groups does the median age exceed the midyear point and in only one group, the nine year girls,

TABLE 1
MEDIAN AGE IN SEX-AGE GROUPS

	7 yrs.		8 yrs		9 yrs		10 yrs	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
Md	7.34	7.46	8.46	8.40	9.37	9.50	10.26	10.33
Q_3	7.61	7.70	8.71	8.73	9.71	9.73	10.63	10.62
Q_1	7.21	7.09	8.17	8.09	9.09	9.21	10.01	10.08
Q	.20	.305	.27	.32	.31	.26	.31	.27
PE_{md}	.035	.054	.047	.057	.055	.046	.055	.048
N	50	50	50	50	50	50	50	50

does it reach it. Likewise, Q_3 at no place exceeds the three quarter of the year mark and Q_1 in each group is below the one quarter of a year mark. The two greatest median age differences occur in years seven and nine. The first of these is .12 of a year, and is only about 1.87 times its probable error. The second is .13 of a year and only about 1.81 times its probable error. Clearly then the two sex groups are so similar chronologically that any comparisons drawn from the standpoint of age should be significant.

Wheat (11, 15) shows in his study of children responses that the chronological age, the grade, and the modal-age-within-the-grade paralleled each other closely and suggests that "when a more extended study of association responses in children is made for the purpose of establishing frequency tables for various developmental levels, it will make little difference whether the tables are set upon the age, the grade, or the modal-age-within-the-grade basis." Table 2 shows the distribution of cases by grades, and in the main, tends to substantiate Wheat's point.

TABLE 2
AGE-GRADE DISTRIBUTION OF CASES

Grade	7 yrs		8 yrs		9 yrs		10 yrs		All Ages		
	B	G	B	G	B	G	B	G	B	G	B&G
6								8	2	8	10
5				1	8	9	19	18	27	28	55
4	1	1	12	15	25	24	15	15	53	55	108
3	2	14	16	22	8	7	8	2	34	45	79
2	32	28	16	9	7	9	4	4	59	50	109
1	15	7	6	3	2	1	2	3	25	14	39
Total	50	50	50	50	50	50	50	50	200	200	400

The children who made up this group of subjects for this experiment came from as widely different environment as could be found in a given section of the country Thirty-eight per cent came from state or private institutions for orphan children, 17 per cent from a private school in a large city, 5 per cent from the public schools of Nashville, Tennessee, and 40 per cent from rural or semi-urban schools

The tabulated data of the reaction times reveal the vast range of differences in reaction time exhibited by individuals in each age-sex group. These differences occurred not only for the reaction times for different words by different individuals but also for the reaction of the same individual for different words in the test One nine year old girl had a range from two seconds on ten words to 244 seconds on the word "young", a seven year old girl had a range from three seconds on 18 words to 241 seconds on the word "doctor"; an eight year old boy had a range from 2 5 seconds on 13 words to 172 seconds on the word "teacher " These, of course, are extreme cases but many individuals had ranges of from 15 to 30 seconds between their lowest and highest reaction times The median reaction times for each stimulus word show a range in the eight year girls' group from 1 38 seconds to the stimulus word *teacher* to four seconds for the word *prison,* and in the boys' seven year group, a range from 93 seconds to the word *milk* to three seconds for the word *flag*

The data which show these differences are so vast and bulky that it is not feasable to present them in the body of this study These differences, therefore, have been studied from the point of central tendencies of time reaction to each stimulus word. The medians of the reaction times for each stimulus word at each age-sex level will be found in Table 3

TABLE 3

Median Reaction Times to Stimulus Words in Terms of Seconds

		7 years		8 years		9 years		10 years		All Ages	
		B	G	B	G	B	G	B	G	B	G
1	money										
	Md	1 51	2 25	1.86	1.86	1 80	2 57	2 00	2 00	1 83	2 13
	PE_{md}	25	29	.30	30	33	20	18	18	117	120
	N	49	49	50	49	49	49	50	48	198	195
2	afraid										
	Md	2 20	2 60	2 94	1 95	2 59	2 58	1 79	2 58	2 40	2 58
	PE_{md}	33	37	38	37	38	37	24	26	170	172
	N	46	46	47	48	45	47	49	47	187	188

TABLE 3 (continued)

	7 years		8 years		9 years		10 years		All Ages	
	B	G	B	G	B	G	B	G	B	G
3. cold										
Md	1.67	1.96	1.85	1.86	1.81	2.66	1.82	2.25	1.82	2.11
PE_{md}	.33	.31	.27	.29	.24	.21	.23	.12	.138	.125
N	49	47	50	50	50	47	50	47	199	191
4. mad										
Md	1.69	2.68	2.60	2.56	2.50	2.92	2.50	2.88	2.50	2.78
PE_{md}	.33	.37	.59	.47	.36	.28	.25	.36	.166	.196
N	46	47	48	47	48	48	50	49	192	191
5. dance										
Md	2.59	3.13	1.94	2.80	2.54	4.56	2.95	3.00	2.57	3.07
PE_{md}	.33	.26	.42	.56	.61	.38	.21	.28	.192	.182
N	45	45	45	46	47	44	48	46	185	181
6. angel										
Md	1.88	2.25	1.77	2.60	2.50	2.96	2.75	2.50	2.19	2.55
PE_{md}	.24	.32	.22	.35	.23	.22	.30	.29	.108	.149
N	47	48	48	48	48	47	49	46	192	189
7. doctor										
Md	1.65	2.38	1.89	1.96	1.90	2.68	2.00	2.50	1.90	2.46
PE_{md}	.32	.27	.32	.30	.28	.26	.27	.20	.151	.139
N	47	50	48	49	38	48	50	48	193	195
8. police										
Md	1.64	2.75	1.60	2.25	1.64	2.61	1.73	2.84	1.64	2.43
PE_{md}	.30	.31	.18	.41	.28	.26	.18	.25	.104	.149
N	49	47	50	46	49	48	50	49	198	188
9. bed										
Md	1.63	1.88	1.81	1.65	1.56	2.00	1.67	2.00	1.62	1.96
PE_{md}	.26	.19	.28	.22	.20	.15	.16	.18	.112	.106
N	50	50	50	50	50	50	50	49	200	199
10. dog										
Md	1.63	1.53	1.81	2.13	1.56	1.96	1.67	1.82	1.65	1.89
PE_{md}	.27	.36	.28	.24	.27	.19	.16	.16	.131	.113
N	48	48	49	49	50	49	50	49	197	195
11. prayer										
Md	1.94	2.90	2.50	2.50	2.54	3.88	2.55	2.88	2.52	2.89
PE_{md}	.34	.27	.40	.36	.26	.32	.34	.36	.153	.173
N	49	49	48	48	49	49	48	48	194	194
12. boy										
Md	1.88	1.58	2.58	1.60	1.67	2.25	1.94	1.41	1.91	1.77
PE_{md}	.35	.29	.46	.38	.27	.28	.29	.18	.128	.14
N	47	48	49	48	48	49	48	49	192	194
13. doll										
Md	1.81	2.64	2.00	2.00	2.25	2.94	2.00	2.67	2.00	2.66
PE_{md}	.39	.36	.39	.37	.23	.16	.20	.25	.138	.151
N	49	49	50	50	50	50	50	50	199	199
14. cheat										
Md	1.94	3.75	2.64	4.53	2.92	3.96	3.71	3.00	2.78	3.36
PE_{md}	.42	.36	.32	.48	.35	.21	.35	.20	.189	.143
N	47	44	44	47	48	49	50	46	189	186

TABLE 3 (continued)

		7 years		8 years		9 years		10 years		All Ages	
		B	G	B	G	B	G	B	G	B	G
15	ghost										
	Md	1.90	2.33	2.38	2.69	2.30	2.90	2.55	2.85	2.34	2.62
	PE_{md}	.29	.28	.35	.41	.27	.24	.28	.24	.140	.145
	N	50	48	49	48	50	49	50	49	199	194
16	prison										
	Md	1.90	4.00	2.38	2.92	2.30	2.98	2.55	2.86	2.34	2.93
	PE_{md}	.31	.52	.36	.49	.29	.24	.18	.25	.144	.166
	N	44	42	47	49	47	47	49	48	187	186
17	dream										
	Md	1.75	3.70	2.72	3.49	2.75	3.13	1.93	2.96	2.33	3.31
	PE_{md}	.33	.45	.39	.55	.39	.28	.22	.36	.182	.181
	N	48	48	47	49	48	45	49	48	192	190
18	flag										
	Md	3.00	2.86	2.92	2.86	2.92	3.50	2.86	2.92	2.92	2.89
	PE_{md}	.41	.39	.52	.40	.38	.24	.29	.20	.193	.158
	N	47	46	48	48	50	50	49	48	194	192
19	girl										
	Md	1.84	2.58	1.92	1.85	2.83	2.08	2.59	2.70	2.26	2.33
	PE_{md}	.31	.38	.32	.26	.36	.08	.24	.20	.162	.139
	N	46	47	50	50	50	49	50	49	196	195
20	dark										
	Md	1.60	2.75	1.88	2.56	1.86	3.41	1.69	2.08	1.78	2.66
	PE_{md}	.23	.18	.31	.29	.20	.21	.18	.16	.125	.091
	N	50	50	49	50	49	50	50	50	198	200
21	papa										
	Md	1.88	2.25	1.91	1.91	1.68	2.25	1.66	1.92	1.73	2.09
	PE_{md}	.47	.28	.40	.28	.30	.20	.18	.10	.176	.092
	N	48	47	49	49	49	48	49	49	195	193
22	friend										
	Md	1.67	2.80	3.13	3.00	2.69	3.25	2.96	3.75	2.93	3.13
	PE_{md}	.39	.29	.47	.48	.50	.38	.28	.22	.207	.163
	N	48	45	47	48	50	48	50	50	195	191
23	fun										
	Md	1.98	2.78	1.85	1.97	2.50	2.60	2.00	2.77	1.96	2.69
	PE_{md}	.29	.22	.22	.37	.35	.15	.32	.28	.135	.135
	N	50	47	48	49	50	48	50	49	198	193
24	fuss										
	Md	1.71	2.67	2.50	2.25	1.73	2.58	2.50	2.53	2.12	2.56
	PE_{md}	.32	.30	.39	.27	.21	.16	.31	.18	.135	.101
	N	49	47	48	49	49	50	50	49	196	195
25	sick										
	Md	1.75	1.97	1.90	2.42	2.19	2.90	1.93	2.75	1.92	2.59
	PE_{md}	.30	.27	.28	.36	.31	.31	.26	.28	.148	.141
	N	49	47	48	49	49	48	49	50	195	194

TABLE 3 (continued)

		7 years		8 years		9 years		10 years		All Ages	
		B	G	B	G	B	G	B	G	B	G
26	run										
	Md	1 68	1 91	1 78	2 61	2 00	2 61	2 25	2 61	1 89	2 26
	PE_{md}	32	42	29	37	37	18	20	17	147	128
	N	49	48	50	49	48	49	50	50	197	196
27	frog										
	Md	1 58	2 53	1 95	1 38	1 75	2 62	2 06	2 06	1 85	2 57
	PE_{md}	21	27	31	47	19	16	13	23	109	123
	N	49	45	50	47	48	48	49	48	196	188
28	game										
	Md	1 75	2 57	1 89	1 77	1 82	3.61	1 91	2 55	1 86	2 56
	PE_{md}	30	21	38	37	30	.31	15	07	.147	115
	N	48	47	46	45	48	46	50	49	192	187
29	milk										
	Md	93	1 83	1 83	1 60	1.51	1 87	1 68	1 89	1 60	1 85
	PE_{md}	14	24	22	20	19	.11	20	13	105	.075
	N	50	50	50	50	49	50	50	50	199	200
30	teacher										
	Md	1 71	1 38	1 67	1 85	1 77	2 75	1 98	2 29	1 74	2 07
	PE_{md}	22	24	28	25	19	.21	20	17	122	109
	N	48	50	48	50	50	50	49	50	195	200
31	sing										
	Md	1 79	2.82	1 98	2 25	2 38	2 79	2 53	2 68	2 18	2 74
	PE_{md}	30	.38	30	32	27	30	.33	24	142	138
	N	48	47	49	49	49	48	49	49	195	193
32	young										
	Md	1 67	1 95	1 91	1 96	1 96	2 97	1 63	2 13	1 79	2.05
	PE_{md}	.32	29	.29	31	.25	26	20	20	139	125
	N	47	46	49	49	50	49	49	47	195	191
33	love										
	Md	1 71	2.60	2 56	2 85	2 00	3 68	2 88	2 75	2 28	2 80
	PE_{md}	.29	36	32	.38	37	22	39	20	.171	153
	N	47	48	48	49	48	50	48	49	191	196
34	star										
	Md	1 71	2 38	1 78	1 89	1 95	1 68	2 13	2 54	1.87	2 13
	PE_{md}	22	19	22	.22	20	19	24	14	106	181
	N	50	50	49	49	50	49	50	50	199	198
35	whip										
	Md	1 61	2 71	2.00	1 94	2 52	2.71	2 13	2 82	2.07	2 71
	PE_{md}	21	45	30	39	26	18	10	19	118	143
	N	49	46	48	50	47	50	47	48	191	194
36	circus										
	Md	1 70	2 61	2 61	2.00	1 59	2 53	1 91	2 94	1.81	2.57
	PE_{md}	.20	26	23	.27	23	.24	.12	32	110	131
	N	50	46	49	50	49	49	50	48	198	193

TABLE 3 (continued)

		7 years		8 years		9 years		10 years		All Ages	
		B	G	B	G	B	G	B	G	B	G
37	coward										
	Md	1.88	3.70	2.67	2.83	2.25	3.61	1.67	3.56	2.11	3.59
	PE_{md}	21	41	28	47	21	41	26	.20	136	190
	N	46	44	46	46	49	47	50	44	191	181
38	mother										
	Md	1.94	1.83	1.97	1.83	1.96	2.59	1.80	2.13	1.95	1.98
	PE_{md}	32	32	31	19	25	24	40	18	137	091
	N	49	49	49	50	49	50	49	50	196	199
39	tunnel										
	Md	1.75	2.58	2.05	2.67	2.75	3.00	2.50	2.63	2.28	2.65
	PE_{md}	30	39	33	41	30	35	19	20	.151	164
	N	48	40	47	48	49	46	50	48	194	182
40	slap										
	Md	1.63	2.67	1.88	2.53	2.50	2.60	2.57	2.83	2.19	2.64
	PE_{md}	.23	32	22	39	30	21	.28	17	133	138
	N	47	48	49	49	50	48	49	49	195	194
41	pretty										
	Md	1.56	2.63	2.38	1.83	1.96	2.50	2.53	2.55	2.17	2.53
	PE_{md}	33	40	37	30	37	21	28	.23	167	138
	N	47	45	49	48	49	48	49	50	194	191
42	help										
	Md	1.71	2.75	2.71	2.89	3.54	3.55	2.96	3.69	2.84	3.22
	PE_{md}	33	45	32	38	44	32	.22	29	150	163
	N	47	45	48	48	47	49	47	47	189	189
43	park										
	Md	1.65	2.00	1.84	2.56	1.96	2.69	1.89	2.68	1.87	2.62
	PE_{md}	23	26	25	37	28	25	16	19	122	127
	N	48	46	47	48	50	49	50	48	195	191
44	hungry										
	Md	1.63	2.55	1.89	2.33	1.88	3.65	1.89	2.89	1.88	2.72
	PE_{md}	23	25	30	31	21	22	18	18	129	113
	N	48	50	49	48	49	50	50	49	196	197
45	Sunday										
	Md	1.60	2.56	1.84	1.90	2.00	2.54	1.82	2.53	1.83	2.53
	PE_{md}	22	36	33	38	22	17	20	.14	107	108
	N	48	48	49	50	50	50	50	50	197	198
46	go										
	Md	2.58	2.86	2.55	2.75	2.53	2.99	1.87	3.13	2.54	2.93
	PE_{md}	32	42	39	38	30	22	19	22	.53	.146
	N	49	48	49	48	49	49	49	47	196	192
47	happy										
	Md	2.10	2.75	2.25	2.89	2.33	2.89	1.78	2.86	2.18	2.87
	PE_{md}	37	26	24	44	31	19	35	17	152	112
	N	48	47	47	49	50	50	49	49	194	195

TABLE 3 (continued)

	7 years		8 years		9 years		10 years		All Ages	
	B	G	B	G	B	G	B	G	B	G
48 mamma										
Md	95	2 00	2 63	1 83	1 78	2 58	1 88	2 25	1 83	2 13
PE_{md}	30	38	39	22	30	20	27	14	147	103
N	47	48	49	48	49	49	49	49	194	194
49 school										
Md	1 97	1 94	1 88	2 08	1 91	2 67	1.83	2 00	1 90	2 04
PE_{md}	25	27	22	30	18	21	18	11	100	114
N	49	50	48	49	50	50	50	50	197	199
50 laugh										
Md	1 80	2 78	2 33	1 85	1 98	2 83	2 33	2 73	2 16	2 76
PE_{md}	33	37	30	21	31	21	.35	.16	156	126
N	48	49	48	48	49	50	50	50	195	197
51 stealing										
Md	1 71	2 70	2 17	2 56	1 94	3 00	1 97	3 00	1 96	2 85
PE_{md}	31	26	.31	46	28	52	26	24	143	157
N	48	38	49	47	46	44	45	42	188	171
52, water										
Md	1 65	1 87	1 79	1 70	1 83	1 92	1.64	2 42	1 72	1 90
PE_{md}	21	25	22	18	20	12	23	26	103	090
N	50	49	49	49	50	50	50	49	199	197
53 home										
Md	1 64	2 71	1 83	1 94	1 86	2 59	1 96	2 67	1 85	2 63
PE_{md}	.23	30	29	30	25	16	19	16	128	096
N	49	46	49	48	50	48	50	50	198	192
54 book										
Md	1.58	1 83	1 78	1 79	1 72	1.96	1 82	2 57	1 75	1 90
PE_{md}	.20	.17	21	26	19	.12	15	106	.097	076
N	50	49	49	49	49	49	50	50	198	197
55 baby										
Md	1.55	1 92	1 85	1.82	1 84	2.65	1 97	2.73	1 85	2 29
PE_{md}	.34	20	22	22	28	.23	27	17	145	071
N	47	48	49	48	47	49	49	50	192	192

The existence of marked differences in reaction between individuals at the same sex and age level are interesting but this discovery merely gives a point of departure for further investigation It is pertinent, therefore, to the present investigation to discover any differences in reaction times which may exist between the age-sex groups and to investigate the extent of these differences. In other words, the present investigator is interested in discovering if the reaction time of boys is shorter than that of girls and if the reaction time at one age level is longer or shorter than at another age level In order to investigate these differences the median reaction time for each age-sex group has been calculated by securing the median of the median reaction times to each of the stimulus words appearing in the test used in this investigation. The results of these calculations are found in Table 4

TABLE 4
MEDIAN OF TIME RESPONSES TO ALL STIMULUS WORDS IN SECONDS

Age	7		8		9		10		All Ages	
Sex	B	G	B	G	B	G	B	G	B	G
Md	1.72	2.56	1.94	2.13	1.97	2.59	1.96	2.65	1.94	2.59
Q_3	1.85	2.74	2.51	2.61	2.46	2.88	2.52	2.86	2.34	2.79
Q_1	1.64	1.96	1.84	1.86	1.81	2.54	1.83	2.43	1.82	2.19
Q	105	39	.34	38	.33	.17	35	22	.26	.30
PE_{md}	016	065	087	.064	085	028	089	037	014	051
N	55	55	55	55	55	55	55	55	55	55

The first significant fact which this table reveals is that in each age level the median time reaction of girls is greater than that of the boys and the lowest median time of the girls, 2.13 at the eight year old level, is greater than the highest median time of the boys, 1.97 at the nine year old level. At each level the difference is statistically reliable as the computation of the probable error of the difference shows The probable error of the difference between boys and girls at the seven year level is .021, between boys and girls at the eight year level, 029, between boys at the nine year level and girls at the eight year level, 027, boys and girls at the ten year level, 026

It is evident then that a significantly longer reaction time for girls than boys is revealed by the free word association test used in this experiment The present investigator is inclined to agree, from his experience in giving the tests, with Hull and Lugoff that the longer reaction time of girls is in part explained by their greater willingness to cooperate and greater desire to make a better record on the test

No consistent difference between the median reaction times of boys and girls viewed chronologically is revealed by the present investigation The only statistically reliable difference in the case of boys is found between the seven year boys and the other age groups There seems to be, then a slight tendency for boys, eight, nine, and ten years old, to give a longer reaction time than do boys seven years old.

The median reaction times of girls viewed chronologically give even a less conclusive picture of difference than do the reaction times of boys. It is noticeable, moreover, that the differences between eight year old girls and the other chronological levels are the only statistically reliable differences found at any age level for girls The reaction time for eight year old girls is always less than the reaction time of the girls at the other age levels but this decided drop in reaction time cannot be explained by the data secured in the present investigation

If, however, the girls' reaction time is compared with boys', omitting the median time reaction at the eight year level, a decided similarity is revealed, that is, a short reaction time for both boys and girls at the seven year old level, a gradual and, for the girls not a statistically reliable, increase in length of reaction time to the nine year old level, and a slight and statistically unreliable decrease in length of reaction time at the ten year old level for both boys and girls.

It must be concluded, however, from the data at hand that the only definite trend shown in differences in reaction between age levels is the tendency for boys eight, nine, and ten years old to give longer reaction times than do boys at the seven year old level.

In so far as chronological age in a normal distribution of cases educationally is indicative of mental age, and Wheat points out that whether the chronological age, the grade, or modal-age-within-the-grade is used in examining children's reactions to a free word association test the results are the same, no evidence such as that suggested by Otis can be found in this investigation that speed of reaction depends upon the degree of mental development of the child A child retarded mentally could do as did seven subjects examined in this experiment, he could answer "I," or a similar word to every stimulus word more rapidly than a mentally normal child could react with logical responses

In order to discover whether children of elementary school age give shorter reaction times to preferred responses than do responses not so classified the following data were calculated The median reaction times have been taken for every tenth stimulus word as the words appear in the test used in this experiment These words are: 10 dog; 20 dark; 30 teacher, 40 slap, 50 laugh The median reaction time for each of these words has been compared with median reaction time for responses to these words which have occurred 20 per cent or more times. In addition, the median of the preferred responses have been compared with median reaction time for each age-sex group. It must be added, however, that although it is felt that sampling of stimulus words is adequate, the conclusions drawn from these data cannot be considered final and must be regarded as merely indicative of certain definite trends. Table 5 gives the median reaction times of the selected stimulus words, the median reaction

TABLE 5

MEDIANS—PREFERRED RESPONSES TO EVERY TENTH STIMULUS WORD, ALL RESPONSES TO EVERY TENTH STIMULUS WORD, AND ALL RESPONSES IN EACH AGE-SEX LEVEL

Medians in Terms of Seconds

	7		8		9		10	
	B	G	B	G	B	G	B	G
Median of All Responses—Age-Sex Levels	1 72	2 56	1 94	2 13	1 97	2 69	1 96	2 65
10 dog								
Md All Resp.	1 63	1 53	1 81	2,13	1 56	1 95	1 67	1 82
Md Pref Resp	1 58	1 50	1 67	1 67	46	1 64	86	90
Q	1 43	1 11	1 22	1 10	74	1 64	86	90
PE_{md}	54	27	48	37	26	73	15	45
N	11	26	10	14	14	14	21	23
20 dark								
Md All Resp	1 60	2 75	1 88	2 56	1 86	3 41	1 69	2 08
Md Pref Resp	1 75	1 00	1 50	2 00	1,55	2 50	1 50	1 75
Q	1 27	82	1 21	1 20	,74	1 00	68	60
PE_{md}	41	24	34	38	22	31	,19	16
N	15	18	20	16	17	16	20	17
30 teacher								
Md All Resp	1 71	1 38	1 67	1 85	1 77	2 75	1 98	2 29
Md Pref Resp	50	1 75	1 00	2 00	41	2 55	2 50	1 78
Q	1 91	67	65	1 50	50	36	1 34	22
PE_{md}	25	22	19	54	,17	13	,53	06
N	33	14	18	12	13	11	10	14
40 slap								
Md All Resp	1 63	2,67	1 88	2 53	2 50	2 60	2 57	2 83
Md Pref Resp	45	2 70	41	1 25	2 67	2 25	2 92	2 65
Q	,182	95	1 27	1 68	1 63	1 00	1 53	,91
PE_{md}	73	34	50	63	59	,38	51	29
N	11	12	10	11	12	11	14	15
50 laugh								
Md All Resp	1 80	2 78	2 33	1 85	1 98	2 83	2 33	2 73
Md Pref Resp	37	1 67	40	1 25	45	3 00	2 25	2 00
Q	28	1 06	10	3 06	,14	65	1 94	1 01
PE_{md}	10	38	36	1 15	52	26	65	37
N	12	12	12	11	11	10	14	12

times of the preferred responses to these words, and the median reaction times of all responses in each age-sex group

An examination of Table 5 reveals that the tendency to give shorter reaction times when the response is a preferred response is not consistent for all the stimulus words examined at each age-sex level. The stimulus word *dog* shows at all age-sex levels a shorter reaction time when the response is a preferred response The stimulus word *dark* shows only one age-sex level with the preferred responses having a larger reaction time than the responses not classified as preferred The other stimulus words of this group show varying tendencies at the different age-sex levels.

The tendency for boys as a whole to give shorter reaction times when responding with a preferred response is not as marked as is the tendency for girls Girls, considered regardless of age differences, show a consistent tendency to give shorter time reactions for preferred response than for responses which do not belong to the preferred category. If, however, both sexes are taken together regardless of age there is a tendency to give shorter reaction time when responding with preferred responses than when responding with reactions not so classified

CONCLUSIONS

The data secured in the present investigation of the reaction times of elementary school children by the free word association method point to the following conclusions·

(1) That differences in reaction times of the subjects of this experiment are greater and more variable than difference in median reaction times of any age or sex groups.

(2) That boys show a consistently shorter median reaction time than do girls.

(3). That age level difference in reaction time show no consistent tendencies It is noted, however, that (*a*) the reaction times of boys tend to increase in a statistically reliable difference from the seventh to the eighth year and remain comparatively constant through the eight, nine, and ten year levels, and further, that (*b*) girls show no statistically reliable difference in reaction time except at the eight year old level which is shorter than at other levels

(4) That children of elementary school age tend to give shorter

time reactions when their verbal reactions are of the "preferred" response variety

(5) That the degree of mental development plays only a small part in the speed of reaction time

(6) That previous degree of familiarity with the stimulus word tends to affect the reaction times of individuals in this experiment

(7) That further investigation, which will secure introspective data from the children who serve as subjects, is necessary before the etiological factors and the significance of differences in reaction time of children can be determined

REFERENCES

1 CASON, H Association in relation to feeling and gross bodily movement *Amer J Psychol*, 1934, **46**, 207-228

2 CATTELL, J McK Experiments on Association of ideas *Mind*, 1887, **41**, 68-74

3 HULL, C R, & LUGOFF, L S Complex signs in diagnostic free association *J. Exper Psychol.*, 1921, **4**, 11-136.

4 JUNG, C G Collected papers on analytical psychology, 2nd edition London Balliere, Tendall & Cox, 1920

5 MILES, C C, & TERMAN, L M Sex differences and association of ideas. *Amer J Psychol*, 1929, **40**, 165-206

6. OTIS, M A study of association in defectives *J Educ Psychol.*, 1915, **6**, 271-288.

7. ROSANOFF, A J Manual of psychiatry New York John Wiley, 1920

8 ROSANOFF, I R, & ROSANOFF, A J A study of association in children. *Psychol Rev*, 1913, **20**, 43-89

9 RUSK, R R Experiments on mental associations in children *Brit J Psychol*, 1909-10, **3**, 349-385

10. WELLS, F L Some properties of the free association time *Psychol. Rev*, 1911, **18**, 1-23

11 WHEAT, L B Free association to common words *Contrib Educ, No. 498, Teach Coll* New York Columbia Univ, 1931

12 WHITE, M N, & POWELL, N The differential reaction time for pleasant and unpleasant words *Amer J. Psychol*, 1936, **13**, 126-133

13 WOODROW, H, & POWELL, F Children's association frequency tables *Psychol Monog*, 1916, **22**, No. 5, Whole No 97

14 WYMAN, J B. Tests of intellectual, social, and activity interests In *Genetic studies of genius*, Stanford Univ Press, 1925

15 WYATT, H G Free word association and sex differences *Amer J Psychol*, 1932, **44**, 454-472

North Carolina State College
Raleigh, North Carolina

SHORT ARTICLES AND NOTES

SOME CHARACTERISTICS OF THE WRITTEN SPEECH OF STUTTERERS I*

JON EISENSON

This paper presents the results of a preliminary study of the written speech of stutterers. We are desirous of learning whether the stutterer carries over into his written speech the characteristics which are manifest in his spoken speech. II Meltzer (1) studied the speech of stuttering children on the basis of verbal responses to the ink blots of the Rorschach Personality Test. He concluded that the speech of the stutterer is characterized by verbal trial and error, and that the stutterer is more talkative than the normal speaker. (The number of words spoken in a given time period was taken as the measure of talkativeness.)

In the present study we chose a fundamentally different type of test situation. Because we believe that the way to study the speech of a stutterer is to examine his verbal utterances in what is for him a normal speech situation, we asked our subjects to write for fifteen minutes on a topic of current interest to them. The topic chosen was "Amalgamation of Extra Curricular Activities at the College." The topic was of no special significance except that it was widely discussed by the students at the college. The subjects were a group of 15 undergraduate male stutterers who attended the college speech clinic, and two control groups of 15 normal speakers matched with the stutterers for age and class in college. The number of words written was taken as the measure of "talkativeness." The number of words crossed out was taken as an indication of "verbal trial and error."

RESULTS

The results of the experiment are indicated in Table 1. We see that

TABLE 1

COMPARISON BETWEEN STUTTERERS AND NON-STUTTERERS IN RESPECT TO NUMBER OF WORDS WRITTEN AND NUMBER OF WORDS CROSSED OUT IN COMPOSITION WRITING

	Stutterer	Non-Stutterer
Mean number of words written	98	195.6
σ of mean of no. of words written	8.37	11.9
Mean per cent of words crossed out	4.7	1.7
σ of mean per cents of no. of words crossed out	2.1	29

---

*Received in the Editorial Office on December 11, 1936.

the mean number of words for the stutterers is half that of the normal speakers, 98 compared with 195 6 The difference between the means of the two groups as measured by the formula $\frac{\text{Difference}}{\sigma\ \text{difference}}$ is 6 7, indicating unquestioned statistical significance The percentage of words crossed out by the stutterers is 4 7 compared with 1 7 for the normal speakers The number of words involved was too small to establish statistical significance

CONCLUSIONS

The findings indicate that in written speech stutterers are less talkative than normal speakers This is not consistent with Meltzer's findings in regard to the spoken speech of the stutterer The differences may be explained by the probability that the stutterer inhibits responses which he does not put on paper Despite this inhibition of response, the stutterer's written speech, on the basis of the percentage of words crossed out, was more marked by trial and error behavior than that of the normal speaker's The results suggest that the stutterer's difficulty in speech may be a manifestation of an inability to sustain coherent thought and to express such thoughts verbally. In a free association situation where sustained coherent thought is not called for, as in Meltzer's experiment, this difficulty is evidently not present

REFERENCE

1 MELTZER, H Talkativeness in stuttering and non-stuttering children *J. Genet Psychol.*, 1935, 46, 371-390

Brooklyn College
Brooklyn, New York

PLANTAR RESPONSES IN INFANTS FOLLOWING A STARTLE STIMULUS*[1]

FRANCES M CLARKE, WILLIAM A. HUNT, AND EDNA B HUNT

An investigation of the startle pattern response in infants following a sudden shot stimulus revealed a wealth of plantar responses (2) Since plantar response is not found in adults as part of the startle pattern, and

---

*Received in the Editorial Office on October 28, 1936

[1] This study was facilitated by a National Research Council grant to one of the authors Thanks are due Miss Mildred McGourty for assistance in all phases of the work Particular gratitude is due Miss Elizabeth Munger, director of the Connecticut State Farm for Women at Niantic, and to Dr. Freitag and all the other members of the staff for their cooperation during the work

since its diverse forms in infants give no indication of the regularity and uniformity typical of the rest of the response, it seems best to treat it separately. This decision is reinforced by the interest of the plantar response per se, and by the further information offered on many of its problems by our data. The mere fact that most of the previous material has been furnished by experiments utilizing cutaneous stimulation makes their elicitation by an auditory stimulus interesting. Our data also offer material bearing on the questions of the appearance of any qualitative regularity in the plantar response, the relationship between age and type of plantar response, the relationship between plantar response and withdrawal flexion of the limb, and the effect of peripheral tone on the response.

PROCEDURE

A total of 63 responses in infants ranging from 8 days to approximately one year in age, are reported. These were selected from the original group of 97 responses (2) solely on the basis of their availability for careful observation. This selection was necessary since some of the older infants were standing, some had their feet outside the photographic field, some were moving at the time, and in some cases the photography was not sufficiently sharp to render fine judgments of toe movement possible. We have also discarded from our table the two cases (one 17 months, one 15 months), who showed no plantar responses. The children were photographed lying on their backs. The stimulus was delivered by an electrically fired 22 caliber revolver in circuit with a time marker. Scoring was done by a committee of three judges who viewed the pictures in slow motion. The technique is reported at length elsewhere (2).

RESULTS

The majority of our cases, (63 out of 65) showed plantar responses to the shot stimuli. The variety and diversity of these are shown in Table 1. In this table the original position of the toes is designated by the large letters. The response to the subsequent shot is designated by small letters. Extension is designated by the letter *E*, flexion by *F*, normal by *N*, ? designates cases where the response was uncertain, and *O* designates no response. Mixed responses are indicated by combinations of small letters in the order in which the responses appeared. That plantar responses can be elicited by a sudden loud stimulus and that such elicitation is usual is shown by our results. Such plantar responses have not appeared in adults (1), but are found in primates (4). As Table 1 shows, these responses include flexion and extension and are not limited to any one type. The Babinski is included among those responses elicited. This diversity of response agrees with the findings of Richards and Irwin (6), and of Pratt, Nelson, and Sun (5), for cutaneous stimulation. No particular type of response such as flexion or

TABLE 1

Plantar Responses in Infants

Age in Days	8	13	14	20	23	27	37	42	48	60	61	63	65	71	74	77	78	101	103	109	111
Right Big Toe	F?	Ne	Ee	Ff	Nf	Ne	Efe	Ee	Nfe	Ne	Nef	Ff	Ff	Nf	Ne	Ffe	Nfe	Nef	Ne	Ne	Nf
Right Little Toes	Ne	Ne	Ee	Ff	Nf	Ne	Nf	Ee	Nf	Ne	Nf	Ff	Ff	Nf	Ne	Fe	Fe	Nf	Ne	Ne	Ff
Left Big Toe	Ne	Nf	Eee	Fe	Nee	Nf-e	Efe	C	Nef	Ne	Nef	Ff	Fe	?f	Ee	Ff	Efe	E?	Ne	Ee	Nf
Left Little Toes	Ne	Nf	Ee	Ne	Ne	Nfe	Nf	C	Nf	Ne	Ne	Ff	Ff	Ff	N?	Ff	Ff	N?	Ee	Ee	Nf-e
Age in Days	113	124	129	132	132	151	159	161	163	168	173	188	205	214	221	225	230	230	233	234	240
Right Big Toe	Nf	Ff	Ef	?	Ef	Nf	Nf	No	Eee	Efe	Ne	Ne	Ne	No	?	Ee	?	Eef	Ne	Ne	?
Right Little Toes	Nf	Ff	Nf	?	Nf	Nf	Nf	No	Nee	Nfn	Nf	Ne	Ne	Nf	?	No	?	Nef	Ne	Ne	?
Left Big Toe	*Ff*	*Nf*	*Ef*	*Nf*	*Ne*	?	*Nf*	*Ee*	*Nfe*	*Ef*	*Nfe*	Ne	Nef	Nfe	Nf	Ee	Fo	E?f	?e	Nf	Nfe
Left Little Toes	Ff	Nf	Ne	Nf	Ne	?	Nf	Ef	Ee	Ne	Nfe	Nf	Ne	Ne	Nf	Ee	No	Ee?	?e	Nf	No
Age in Days	244	250	250	253	263	270	278	279	279	284	284	286	304	305	307	327	340	355	360	368	378
Right Big Toe	F?	Ne	En	?	Nf	Nfe	Nf	?	Nfe	Nf	Ft	Nf	Nf	Nfe	No	Nf	Fen	Ff	?	No	Nf
Right Little Toes	Ne	Nf	Nf	?	N?	Nf	Nef	?	Nf	Nf	Ff	Nfe	Nf	Nf	Nf	Nf	Fen	Ef	?	E?	Nf
Left Big Toe	Nf	Eo	Ff	Nf	C	Nfe	Ene	Ffn	Ne	Nf	Nfe	Ff	C	Nfe	No	Nf	Fnf	Nf	Ene	Ff	Ne
Left Little Toes	Ne	N?	No	Nf	C	Nfe	Nfe	Ffn	Nfe	Nf	Nfe	Ff	C	Nf	Nf	Nf	Fnf	Nf	Ene	Nf	Ne

Key-
N or n—normal
F or f—flexion
E or e—extension
?—doubtful
o—no response
Large letters indicate posture before shot, small letters posture afterward

extension seems to stand out sufficiently to warrant its being regarded as the normal physiological function at any age level within the range of our subjects. While the infantile Babinski response in both feet is found in our youngest case (8 days), both flexion and extension are found from 13 days on. The only difference with respect to age level is that the younger the infant the more violent the reaction.

There is no confirmation for the view that extension of the toes is associated with the withdrawal flexion of the limb. This withdrawal flexion of the limb is typical of the startle pattern, but it is associated in our results with both flexion and extension of the toes.

The suggestion made by Bourguignon (1) and more recently by Juarros (4), that peripheral tone may be a factor in determining whether the plantar response is one of flexion or extension is not clearly supported by our data. Rather than showing that previous extension predisposes to flexion in the subsequent plantar response, and that previous flexion predisposes to extension, our results show the opposite. Previous flexion seems to encourage a more complete flexion in the subsequent plantar response in the majority of cases. For previous extension, the matter is less clear due to the inadequate number of cases. We can conclude, however, that the relation of peripheral tone to plantar response is not as simple as has been thought heretofore.

SUMMARY

A sudden sound stimulus was found to release plantar responses in a group of 63 infants from eight days to one year in age. These responses are of varied types and show little relationship to such factors as age, limb withdrawal, and peripheral tone.

REFERENCES

1. BOURGUIGNON, G. Les conditions périphériques du réflex plantaire normal et du signe de Babinski. *Rev. Neur.*, 1927, **1**, 1081-1085.
2. HUNT, W. A., CLARKE, F. M., & HUNT, E. B. Studies of the startle pattern: IV. Infants. *J. of Psychol.*, 1936, **2**, 339-352.
3. HUNT, W. A., LANDIS, C., & JACOBSEN, C. Studies of the startle pattern: V. Apes and monkeys. *J. of Psychol.*, 1937, **3**, 339-344.
4. JUARROS, C. Aportacion al conocimiento de algunos problemas planteados por el estudio del signo de Babinski. *Siglo Méd.*, 1930, **86**, 689-694.
5. PRATT, K. C., NELSON, A. K., & SUN, K. H. The behavior of the newborn infant. Columbus: Ohio State Univ. Press, 1930.
6. RICHARDS, T. W., & IRWIN, O. C. Studies in infant behavior: II. Plantar responses of infants and young children. Univ. Iowa Studies, 1934.

Connecticut College
New London, Connecticut

APPARATUS

THE TYPEWRITER KEY AS A DEVICE FOR REGISTERING CHOICE TIMING*

MARIE L H FORBES

Figure 1 shows a line of typewritten marks each representing an interval of time, the letter x below the line representing a choice registered at the fifteenth movement of the spacing mechanism, which is not operated by the subject

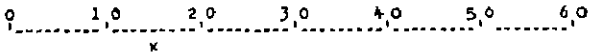


FIGURE 1

Figure 2 shows a cardboard frame for the display of a caption above a blank key (1) (2), the frame being held upright behind the lowest row of keys by means of a crosspiece which passes under the keyboard

The writer gratefully acknowledges the examination of the materials by

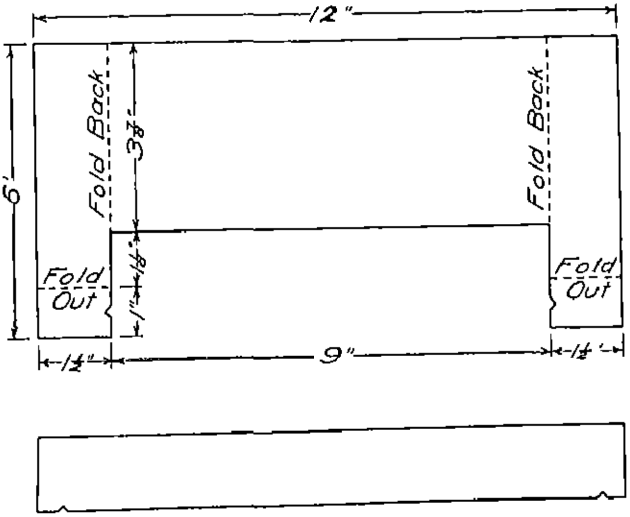


FIGURE 2

---

*Received in the Editorial Office on October 15, 1936

her former instructor, Dr Walter F Dearborn, Director of the Psycho-Educational Clinic, Harvard University

REFERENCES

1. FORBES, M L. H A combination of horn-book and typewriter. *J Genet Psychol*, 1936, 48, 250-252
2. ———. The typewriter key as a device for registering choice *J Genet Psychol*, 1936, 48, 253

8½ Ash Street Place
Cambridge, Massachusetts

BOOKS

KOHTS, N *Infant Ape and Human Child, Instincts, Emotions, Play, Habits* [Ditia chimpanze i ditia tcheloveka, v ikh instinktakh, emotiakh, igrakh, privytckkakh i vyiazitelnykh dvizheniakh] Scientific Memoirs of the Museum Darwinianum in Moscow, v III, 1935 Pp xvi+596 c 145 plates.

Out of Russia comes a striking volume with an arresting title *Infant Ape and Human Child* It appears as a bulky scientific memoir, embellished with 145 plates, reproducing 604 individual photographs A summary in English and the wealth of photographs and numerous drawings, made by the author or under her supervision, render the work useful and highly suggestive to those who cannot read the original Russian

Madam Kohts is unquestionably an observer of ingenuity, patience, and scientific restraint In a double sense this is a contribution to comparative psychology Her basic observations were made from 1913 to 1916 on a male chimpanzee, Joni, during the age period from 1½ to 4 years In 1925 her son Roody was born, she made behavior records of his development from the moment of birth up to the age of 4 years in 1929. The respective infrahuman and human data were therefore separated by a period of twelve years Such a separation of time doubtless reduced the detail and incisiveness of the comparative data, but it also served to temper the final conclusions and to obviate serious preconceptions The author has at least avoided the temptation of magnifying the human aspect of chimpanzee and she brings into clear relief the distinctive characters in the superiority of *homo sapiens* It seems to us that she has succeeded better in this task of critical comparison of two species than have those studies which emphasize somewhat too strongly similarities, identities, and simian precocities.

Madam Kohts has taken fully five years to assemble and to evaluate her comparative findings The text divides into 3 parts I The Behavior of the Infant Chimpanzee, II The Behavior of the Human Child, III Analysis of Behavior of Man & Ape (pages 454 to 488) Chimpanzee earns a special chapter on *Deceptive & Sly Behavior* For Roody there is a special chapter on *Memory & Habit Forming* (conditioned-reflex behavior). There are two equivalent chapters on bodily features, face, arms, hands, legs, static and dynamic postures The line drawings of postural and mimetic attitudes in chimpanzee, are unusually well done One wishes that here the comparison had been carried over more systematically to the human The author finds that the manifestation of emotions as expressed in mimics, pantomime and vocalization is "incomparably more prominent and accentuated in the ape."

In the matter of articulate sounds, however, no comparison whatever can be drawn. All told Joni was capable of emitting when moved by emotion 25 distinguishable sounds. Nearly all of these sounds were also in Roody's repertoire but at the age of one year he was already freely indulging in spontaneous word building.

Although Kohts did not, after the manner of Kellogg, attempt to create an identical environment for the early development of ape and child, she observed her subjects in similar situations. Both "children" were easily taught to soap their hands, to use a handkerchief, and to wipe their mouths with a serviette. The reactions of both to their own mirror image were observed to be strikingly alike, including spying, smiling, touching, grimacing, spitting, menacing gesticulation, and labial vocalization.

Numerous and many sided similarities were discovered in many fields of behavior. The only specific trait of the chimpanzee with apparently no analogy whatever in human behavior was found to be a "modulated grunting *oh-oh* sound ending in a bark and accompanying general excitement." Madam Kohts is also impressed by the fact that she failed to find in the chimpanzee even "a single psychic trait which would not at the same time belong to man at this or another stage of development." But when she passes from analysis to synthesis and asks, "What after all, is the contemporary chimpanzee?" she answers, "Not only is it impossible to say that he is 'almost human,' we must go even further and state quite definitely that he is 'by no means human,'"—and this on the ground of evidence which she sums up as follows:

> (1) In the functional biological field the chimpanzee totally ignores the possibility of walking erect and of freeing his hands for carrying weights, (2) In the sphere of imitation the chimpanzee is devoid of imitation insofar as human sounds are concerned and generally fails to extend or improve his imitatory behavior, (3) In respect of emotional altruistic and social behavior, the chimpanzee fails to understand the advantages of friendly sympathetic intercourse with creatures standing on a lower biological level than himself, (4) With regard to habit-forming the chimpanzee does not improve in the motor habits connected with the use of tools and household implements, (5) In the sphere of playful behavior he does not indulge in creative constructional play.

Yerkes has called attention to the wide gamut of individual differences among chimpanzee. Such differences must be taken into account before generalizing the comparisons arrived at in the present volume. It must also be remembered that the author had no opportunity to observe in her ape the interesting period of development between birth and the age of 1½ years From the standpoint of comparative developmental psychology this is precisely the period which needs parallel study in ape and man Such a study would be less concerned with a comparison of specific abilities than with the deeper biological problem of species differences in developmental mechanics From the standpoint of life cycle, what are the early differences in the ontogenesis of behavior,—in the patterns of behavior growth? This is the fundamental genetic problem

In 1923 Mrs. Kohts published a volume on chimpanzee intelligence ("Untersuchungen uber die Erkenntnisfahigkeiten des Schimpansen," Moscow, pp 453) She is engaged in a third experimental research on the "ability of the chimpanzee to distinguish form, size, quantity, and number and his capability towards ascertaining likeness and dissimilarity, analysis and synthesis" Through this experimental approach she hopes to determine whether chimpanzee harbors those human-like characters which *proanthropus* had to possess in order to become *homo sapiens.*

The present volume therefore lies between two scientific undertakings. In a stimulating way it suggests many vistas of investigation The extensive collection of photographs, sumptuously reproduced on special paper at the end of the book, greatly enhance its value The photographs depict Joni and Roody in naturalistic situations. The expressional behavior of chimpanzee has been caught in great variety and in detail These particular pictures are little short of classic and would have delighted Darwin, whose name adorns the museum under whose auspices this pioneer comparative study of a pan-child and man-child was carried through

ARNOLD GESELL

Clinic of Child Development
Yale University
New Haven, Connecticut

Books Received

Adler, M J Art and prudence. New York Longmans, Green, 1937 Pp xv+686

Castiello, J A humane psychology of education New York: Sheed & Ward, 1936 Pp xxiii+254

Dodge commemorative number of the *Psychological Monographs* Edited by Walter R Miles *Psychol Monog*, 1936, 47 Pp xxxvi+415

Fisher, V E Auto-correctivism Caldwell, Idaho Caxton Printers, 1937 Pp. 337

Iovetz-Tereshchenko, N M Friendship-love in adolescence London Allen & Unwin, 1936 Pp xvi+367

Hardy, M. C, & Hoefer, C. H Healthy growth Chicago Univ of Chicago Press. (Undated) Pp. xii+360

Kirkman, F. B Bird behavior London' Nelson & Sons, 1937. Pp xv+232

Lundholm, H The psychology of belief. Durham, North Carolina Duke Univ Press, 1936 Pp. vii+243

McAllister, A H. Clinical studies in speech therapy. London. Univ London Press, 1937 Pp. xxiv+376

Odenweller, A L. Predicting the quality of teaching New York. Teachers College, Columbia Univ., 1936 Pp. xi+158.

Richmond, W V Personality. New York. Farrar & Rinehart, 1937 Pp xvii+279.

Skinner, C. E. Readings in educational psychology. New York Farrar & Rinehart, 1937 Pp viii+630.

Club Rate For Five International Psychological Journals

	PER ANNUM
THE PEDAGOGICAL SEMINARY AND JOURNAL OF GENETIC PSYCHOLOGY (Quarterly, two volumes per year at $7.00 per volume)	$14.00
GENETIC PSYCHOLOGY MONOGRAPHS (Quarterly, one volume per year)	$7.00
THE JOURNAL OF GENERAL PSYCHOLOGY (Quarterly, two volumes per year at $7.00 per volume)	$14.00
THE JOURNAL OF SOCIAL PSYCHOLOGY (Quarterly, one volume per year)	$ 7.00
THE JOURNAL OF PSYCHOLOGY (Quarterly, two volumes per year at $7.00 per volume)	$14.00
TOTAL	$56.00
CLUB RATE	$50.00

THE JOURNAL PRESS
2 Commercial Street
Provincetown, Massachusetts
U.S.A.